Electron. J. Probab. **19** (2014), no. 83, 1–36.
ISSN: 1083-6489 DOI: 10.1214/EJP.v19-2458

Rank deficiency in sparse random GF$[2]$ matrices

R.W.R. Darling* Mathew D. Penrose† Andrew R. Wade‡
Sandy L. Zabell§

Abstract

Let M be a random $m \times n$ matrix with binary entries and i.i.d. rows. The weight (i.e., number of ones) of a row has a specified probability distribution, with the row chosen uniformly at random given its weight. Let $\mathcal{N}(n,m)$ denote the number of left null vectors in $\{0,1\}^m$ for M (including the zero vector), where addition is mod 2. We take $n, m \to \infty$, with $m/n \to \alpha > 0$, while the weight distribution converges weakly to that of a random variable W on $\{3,4,5,\ldots\}$. Identifying M with a hypergraph on n vertices, we define the *2-core* of M as the terminal state of an iterative algorithm that deletes every row incident to a column of degree 1.

We identify two thresholds α^* and $\underline{\alpha}$, and describe them analytically in terms of the distribution of W. Threshold α^* marks the infimum of values of α at which $n^{-1}\log \mathbb{E}[\mathcal{N}(n,m)]$ converges to a positive limit, while $\underline{\alpha}$ marks the infimum of values of α at which there is a 2-core of non-negligible size compared to n having more rows than non-empty columns.

We have $1/2 \le \alpha^* \le \underline{\alpha} \le 1$, and typically these inequalities are strict; for example when $W = 3$ almost surely, $\alpha^* \approx 0.8895$ and $\underline{\alpha} \approx 0.9179$. The threshold of values of α for which $\mathcal{N}(n,m) \ge 2$ in probability lies in $[\alpha^*, \underline{\alpha}]$ and is conjectured to equal $\underline{\alpha}$. The random row-weight setting gives rise to interesting new phenomena not present in the case of non-random W that has been the focus of previous work.

Keywords: Random sparse matrix; null vector; hypercycle; random allocation; XORSAT; phase transition; hypergraph core; large deviations; Ehrenfest model.
AMS MSC 2010: Primary 60C05, Secondary 05C65; 05C80; 15B52; 60B20; 60F10.
Submitted to EJP on November 23, 2012, final version accepted on September 4, 2014.

1 Introduction

Suppose that $M := M(n,m)$ is an $m \times n$ matrix with entries in $\{0,1\}$, each of whose rows contains at least one 1, for which we seek a left null vector over GF$[2]$, i.e. a row vector $a \in \{0,1\}^m$ such that $aM \equiv \mathbf{0}$ (mod 2), where here and elsewhere $\mathbf{0}$ is the all-0 vector. We are interested in the case where M is sparse and random, as specified below.

*Mathematics Research Group, National Security Agency, USA. E-mail: rwdarli@nsa.gov
†Department of Mathematical Sciences, University of Bath, UK. E-mail: M.D.Penrose@bath.ac.uk
‡Department of Mathematical Sciences, Durham University, UK. E-mail: andrew.wade@durham.ac.uk
§Mathematics Department, Northwestern University, USA. E-mail: zabell@math.northwestern.edu

Let $X_1, X_2, \ldots, X_m$ denote the vectors constituting the rows of M, and let $\sigma(n,m)$ denote the co-rank over GF[2], namely, with 'span' denoting the linear span over GF[2],

$$\sigma(n,m) := m - \dim \operatorname{span}\{X_1, X_2, \ldots, X_m\}. \tag{1.1}$$

Then the number of null vectors of M, including the zero vector, is

$$\mathcal{N}(n,m) = 2^{\sigma(n,m)}, \tag{1.2}$$

which counts the number of distinct solutions in $\{0,1\}^m$, including the zero solution, to

$$a_1 X_1 + \cdots + a_m X_m \equiv \mathbf{0} \pmod 2. \tag{1.3}$$

Note that for a fixed n and a given realization of the sequence of rows $X_1, X_2, \ldots$, the numbers $\mathcal{N}(n,m)$ are nondecreasing as m increases.

Suppose that $n, m \to \infty$, with $m/n \to \alpha > 0$. We study asymptotics of the expected size $\mathbb{E}[\mathcal{N}(n,m)]$ and probability $\mathbb{P}[\mathcal{N}(n,m) > 1]$ of non-triviality of the left null space of $M(n,m)$, in terms of the asymptotic aspect ratio α. In particular, we derive computable thresholds for α at which phase transitions occur. We describe the relevance of the 2-core construction to this question. We also study the rate of exponential decay of the probability that $\mathbf{1} := (1, 1, \ldots, 1)$ is a null vector.

In our probabilistic setting, the rows $X_1, X_2, \ldots$ are independent and identically distributed (i.i.d.) with the law of a random vector $X = X(n) \in \{0,1\}^n$. Our focus is the (very) sparse regime in which the number of non-zero components of X converges in law as $n \to \infty$ to some given *weight distribution*. The existing literature focuses on the simplest case, where the weight distribution degenerates to some constant r.

Before describing our model in detail and presenting our main results (in Section 2), we make some remarks on motivation. Note that $\sigma(n,m) = 0$ if and only if M has row rank m, which occurs if and only if M has column rank m. Thus the absence of non-trivial left null vectors is equivalent to all column vectors in $\{0,1\}^m$ being expressible as a linear combination of the columns of M (with addition modulo 2), or in other words, to there being a solution $x \in \{0,1\}^n$ to $Mx \equiv y$ for all column vectors $y \in \{0,1\}^m$. In the special case of $r = 2$, motivation for considering this question is discussed at the start of [16, Chapter 3]. The following interpretations help to motivate the general case.

A scheduling problem. A tennis club is organizing its annual schedule. There are n playing days, and m potential players. Each player wants to play on a given subset of the days; if there is not a match available on every one of these days, they refuse to pay the annual membership. So that nobody is left out, an even number of players is needed on each day. Each possible schedule satisfying these requirements is a left null vector mod 2; the one with the most units achieves the maximal income for the club.

Randomized Lights Out. This is a variant of the game 'Lights Out' [22]. Each of m lamps can be either on or off, and there are n switches, each of which is incident to a subset of the lamps specified by the matrix M; Lamp i and Switch j are mutually incident if and only if the (i,j) entry of M is 1. If a switch is toggled, the status of every incident lamp is changed (from on to off or off to on). All configurations of on and off lamps are accessible from the 'all off' state by using some sequence of switches if and only if the column rank of M is m.

XORSAT. This is a variant of the random satisfiability problem [21], where there are n Boolean variables which may be deemed true or false. Each row of M represents a clause built as the logical XOR (exclusive OR) involving those Boolean variables corresponding to columns incident to this row, so the clause is true if an odd number of the

variables incident to the row are deemed true. Given a vector $y \in \{0,1\}^m$, finding a solution x to $Mx \equiv y$ corresponds to finding a truth-assignment for the Boolean variables so that each clause i is true if $y_i = 1$ and false if $y_i = 0$. Thus the column rank is m if and only if the problem is satisfiable for all possible choices of y.

A spin-glass model. The relationship between satisfiability problems and spin glasses has already been noted in [21]. In the present instance, consider the following variant of the well-known Sherrington–Kirkpatrick mean-field spin-glass model (see e.g. [23]). There is a random collection of hyperedges on n vertices, represented by the m rows of M. Each hyperedge i has a sign g_i, taking value $(-1)^{y_i}$. Each vertex j is assigned a spin $\sigma_j \in \{-1, 1\}$. The (zero temperature) probability measure on the state-space is concentrated on states of minimal energy, i.e. with maximal value of $\sum g_i e_i$, where e_i is the product of spins at vertices in hyperedge i. The existence of a configuration with all terms in the sum equal to $+1$ is equivalent to the existence of a solution to $Mx \equiv y$.

The Ehrenfest urn and random walk on the hypercube. In the Ehrenfest model of heat exchange (see e.g. [12, p. 121] or [20, §3.5]), a box contains n particles, each either red or blue. At each step, a particle is sampled uniformly from the box and changes its colour. In the case where X has a single unit entry, we may view each row of M as selecting which particle is to be changed at that step. Then $\mathbf{1}$ is a null vector for M if and only if the box returns to the initial state after m steps. This may be phrased in terms of a random walk on a discrete hypercube $\{0,1\}^n$: the event that $\mathbf{1}$ is null corresponds to the walker being back in the initial state after m steps.

The general case, allowing other weight distributions, corresponds to a generalization of the Ehrenfest model whereby multiple 'diffusions' are allowed, i.e. at each step several particles may change colour at once; cf [20, Chapter 10]. This can be similarly interpreted in terms of a walk on a version of the hypercube with additional edges.

There is a large body of work on random matrices and random linear equations over finite fields, including the surveys [16, Chapter 3] and [18, 19]. Problems may also be formulated in terms of *random hypergraphs*: each row represents a hyperedge, and each column represents a vertex (see Section 2.3 below). Generally, such models can be described in the framework of *random allocation* or *occupancy* problems [16, 20, 17].

The null-vector problem in the fixed row-weight case has received several treatments in the literature. It is not easy to reconcile all the existing results, due to differences both in presentation and in the underlying probabilistic models. The present paper provides clarification, including a rigorous justification that the results are unchanged under small perturbations of the underlying model. Our main contribution, however, is the treatment of genuinely *random* row weights, which is new. We mention recent renewed interest in this area in several scientific communities: Alamino and Saad [1] give a statistical physics approach to the null-vector problem; Ibrahimi *et al.* [13] treat the related random XORSAT problem; Costello and Vu [6] study the rank of random symmetric matrices.

Throughout the paper, we extend the function $x \mapsto x^x$, $x > 0$, continuously to $x = 0$, so that $0^0 := 1$. We define the *weight* of a vector $v = (v_1, \ldots, v_n) \in \{0,1\}^n$ to be $w(v) := \sum_{i=1}^n v_i$, i.e., the number of unit entries. For $n \in \mathbb{N} := \{1, 2, \ldots\}$ we write $[n] := \{1, 2, \ldots, n\}$. We write $\xrightarrow{d}$ for convergence in distribution.

2 Results and discussion

2.1 Description of the random matrix model

Given $n \in \mathbb{N}$, suppose that $X = X(n) \in \{0,1\}^n$ is a random row vector, selected according to a probability law of the following form. Let W be an $\mathbb{N}$-valued random

variable ($\mathbb{P}[W \geq 1] = 1$) whose law will be the (limiting) *weight distribution* of X. Let $W_1, W_2, \ldots$ be a sequence of random variables with $W_n \in [n]$ such that $W_n \xrightarrow{d} W$ as $n \to \infty$. Let $w(X)$ have the distribution of W_n, and for each $k \in [n]$ let the conditional distribution of X, given $w(X) = k$, be uniform over $\{x \in \{0,1\}^n : w(x) = k\}$.

Consider i.i.d. random vectors $X_1, X_2, \ldots$ with the same law as X. Let $M := M(n,m)$ be the $m \times n$ matrix whose rows are $X_1, X_2, \ldots, X_m$. Let $\rho(s) := \mathbb{E}[s^W]$ and $\rho_n(s) := \mathbb{E}[s^{W_n}]$ denote the probability generating functions of W and W_n, respectively. We use $\mathbb{P}_{\rho_n}$ and $\mathbb{E}_{\rho_n}$ for probability and expectation for the random matrix model with n columns and row-weight distribution specified by ρ_n. We say W_n are *uniformly bounded* if there is a finite constant r_1 such that $\mathbb{P}[W_n \leq r_1] = 1$ for all n (so $\mathbb{P}[W \leq r_1] = 1$ too).

2.2 Threshold results in the general setting

Given the probability generating function ρ, define the threshold

$$\alpha^*_\rho := \inf\{\alpha \geq 0 : F_\rho(\alpha) > 0\}, \tag{2.1}$$

where we set

$$F_\rho(\alpha) := \log \sup_{\gamma \in [0,1/2]} \left(\frac{(1+\rho(1-2\gamma))^\alpha}{2\gamma^\gamma (1-\gamma)^{1-\gamma}} \right), \quad \alpha \geq 0. \tag{2.2}$$

We note (i) $F_\rho(\,\cdot\,)$ is continuous and nondecreasing, with $F_\rho(\alpha) = 0$ for $0 \leq \alpha \leq \alpha^*_\rho$ but $F_\rho(\alpha) > 0$ for $\alpha > \alpha^*_\rho$; and (ii) if $\mathbb{P}[W \geq 2] = 1$ and $\mathbb{E}[W] < \infty$, then $\alpha^*_\rho \in [1/2, 1)$. These facts, and others, are proved in Lemma 4.1 below. We now present our first main result, describing the threshold behaviour of the expected number of null vectors $\mathcal{N}(n, m_n)$.

Theorem 2.1. *Suppose that $m_n/n \to \alpha \in (0,\infty)$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} n^{-1} \log \mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)] = F_\rho(\alpha). \tag{2.3}$$

Moreover, if in addition there exist $r_0 \geq 3$ and $r_1 < \infty$ such that $\mathbb{P}[r_0 \leq W_n \leq r_1] = 1$ for all n, and $\alpha \in (0, \alpha^*_\rho)$, then as $n \to \infty$,

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)] = 1 + O(n^{2-r_0}). \tag{2.4}$$

We give the proof of Theorem 2.1 in Section 4.5. A key role in our proof is played by the event $A(n,m)$ that the row vector $\mathbf{1} = (1,\ldots,1)$ is null for M, i.e.,

$$A(n,m) := \{X_1 + \cdots + X_m \equiv \mathbf{0} \pmod 2\}. \tag{2.5}$$

Observe that $\mathcal{N}(n,m)$ is the number of collections of rows of $M(n,m)$ which sum to $\mathbf{0}$ (mod 2), and for each set of ℓ rows the probability that it sums to $\mathbf{0}$ is $\mathbb{P}_{\rho_n}[A(n,\ell)]$. So

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)] = \sum_{\ell=0}^{m} \binom{m}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)]. \tag{2.6}$$

The first step in our analysis is to study the asymptotics of $\mathbb{P}_{\rho_n}[A(n,m)]$, which is of its own interest in the context of random allocations; see Section 2.4. The starting point for this analysis is the novel exact formula (3.1) below for this probability in the special *binomial* model, which we show serves as a good approximation for the general case (details are in Section 3). The asymptotic analysis of (2.6), leading to the proof of Theorem 2.1, needs additional work; in particular, 'low weight' null vectors must be dealt with separately (details are in Section 4). Indeed, for $\alpha < \alpha^*_\rho$, the expectation in (2.4) is dominated by null vectors with only 2 non-zeros.

For a fixed n, the number of rows m at which the first non-zero null vector appears,

$$T_n := \min\{m \in \mathbb{N} : X_m \in \operatorname{span}\{X_1, X_2, \ldots, X_{m-1}\}\}, \tag{2.7}$$

is another random variable of interest. Standard linear algebra implies that $T_n \leq n+1$.

We define another threshold, $\underline{\alpha}_\rho$, through an analytic description that needs more notation; the probabilistic interpretation of $\underline{\alpha}_\rho$ is in terms of the *2-core* of $M(n,m)$, as we describe in Section 2.3. For $x \in (0,1)$ set

$$\psi(x) := x + \left(1 + \frac{\rho(x)}{\rho'(x)} - x\right)\log(1-x); \tag{2.8}$$

$$h(x) := -\frac{\log(1-x)}{\rho'(x)}. \tag{2.9}$$

Provided $\mathbb{P}[W \geq 1] = 1$ we can and do extend ψ continuously to $\psi(0) := 0$, since $\rho(s)/\rho'(s) = O(s)$ as $s \downarrow 0$. Note that $h(x) \to \infty$ as $x \downarrow 0$ provided $\mathbb{P}[W \geq 3] = 1$, and that if $\mathbb{E}[W] < \infty$ then as $x \uparrow 1$ we have $h(x) \to \infty$ and $\psi(x) \to -\infty$. Set

$$\alpha_\rho^\sharp := \inf_{x \in (0,1)} h(x), \tag{2.10}$$

and note that $\alpha_\rho^\sharp \rho'(x) \leq -\log(1-x)$ for all $x \in (0,1)$, so integrating from 0 to 1 we get $\alpha_\rho^\sharp \leq 1$, provided $\mathbb{P}[W \geq 1] = 1$. Define for $\alpha \geq 0$, with the convention $\sup \emptyset = 0$,

$$g^*(\alpha) := \sup\{x \in (0,1) : h(x) \leq \alpha\}. \tag{2.11}$$

Observe that if h has unbounded range (e.g. if $\mathbb{P}[W \geq 3] = 1$) then $h \circ g^*$ is the identity map on $[\alpha_\rho^\sharp, \infty)$. See Figure 1 for an example. Define

$$\underline{\alpha}_\rho := \inf\{\alpha > \alpha_\rho^\sharp : \psi(g^*(\alpha)) < 0\}. \tag{2.12}$$

In (2.12), the set defining $\underline{\alpha}_\rho$ is non-empty provided $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W] < \infty$, since as $\alpha \to \infty$ we have $g^*(\alpha) \to 1$ and $\psi(g^*(\alpha)) \to -\infty$.

The relevance of $\underline{\alpha}_\rho$ for the null vector problem is shown by the next result.

Theorem 2.2. *Suppose W_n are uniformly bounded and $\mathbb{P}[W_n \geq 3] = 1$ for all n. Then $\alpha_\rho^* \leq \underline{\alpha}_\rho \leq 1$, and for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}_{\rho_n}\left[(\alpha_\rho^* - \varepsilon)n \leq T_n \leq (\underline{\alpha}_\rho + \varepsilon)n\right] = 1. \tag{2.13}$$

Theorem 2.2 is proved in Section 5.4. The case where $\mathbb{P}[W \in \{1,2\}] > 0$ must be excluded from the statement of Theorem 2.2; different phenomena occur in that case, which is discussed in [7], where the functions ψ and h also play a role.

Usually, we have $\psi(g^*(\underline{\alpha}_\rho)) = 0$. In many cases, ψ has a single zero in $(0,1)$, x_ρ^* say, and $\underline{\alpha}_\rho = h(x_\rho^*)$. But the situation may be more complicated; and $\alpha \mapsto g^*(\alpha)$ typically has at least one discontinuity. We return to this briefly in Section 2.3, and defer a more detailed discussion of the properties of the functions ψ, h, and g^*, and the corresponding thresholds, to Section 5.4 below. Figure 1 provides an example.

We believe that the upper bound in Theorem 2.2 (i.e., $\underline{\alpha}_\rho$) is sharp:

Conjecture 2.3. *If W_n are uniformly bounded and $\mathbb{P}[W_n \geq 3] = 1$ for all n, then T_n/n converges in probability to $\underline{\alpha}_\rho$ as $n \to \infty$.*

An equivalent statement to the fixed-weight case $W = r \geq 3$ of this conjecture seems to have been established recently in the random-XORSAT literature: see Section 2.6.2.

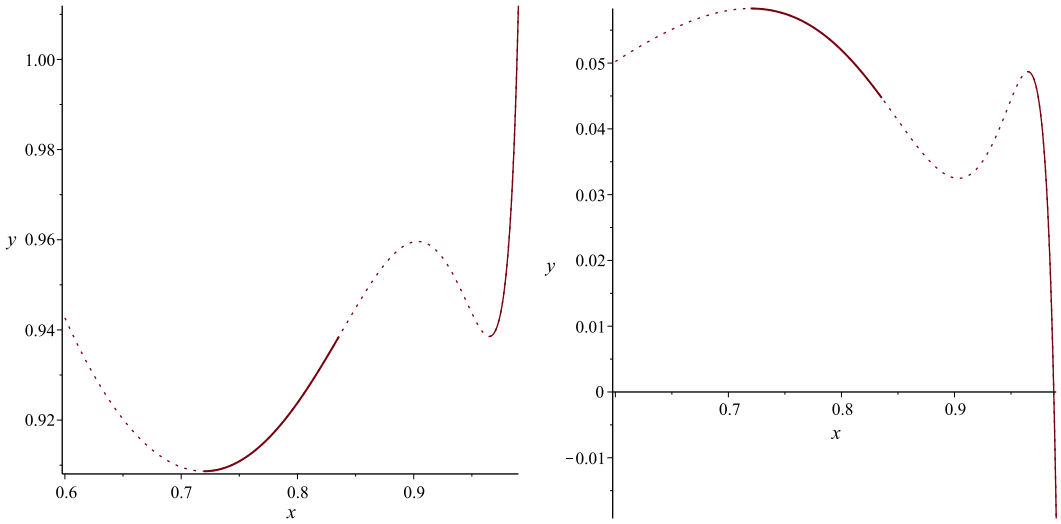

Figure 1: Example with $\rho(s) = 0.9s^3 + 0.1s^{24}$. The left plot shows parts of the curves $y = h(x)$ (all the line) and $x = g^*(y)$ (solid line). The right plot shows parts of the curves $y = \psi(x)$ (all the line) and the locus of $(g^*(\alpha), \psi(g^*(\alpha)))$ (solid line). The left plot shows that $g^*(\alpha)$ has two discontinuities, one at $\alpha = \alpha_\rho^\sharp \approx 0.908654$ and one at $\alpha \approx 0.938536$, with the first corresponding to a jump from $g^* = 0$ to $g^* \approx 0.719682$ and the second to a jump from $g^* \approx 0.835696$ to $g^* \approx 0.964919$. The right plot shows the single positive solution of $\psi(x) = 0$ at $x = x_\rho^* \approx 0.987817$, so $\underline{\alpha}_\rho = h(x_\rho^*) \approx 0.991613$. It is not a coincidence that the curves h and ψ seem to mirror each other: see Lemma 5.8 below.

2.3 2-cores and random hypergraphs

To describe the probabilistic interpretation of $\alpha_\rho^\sharp$ and $\underline{\alpha}_\rho$ we need additional terminology. Given a set $\mathcal{V} = \{v_1, \ldots, v_n\}$, whose elements we call *vertices*, a non-empty subset of $\mathcal{V}$ is called a *hyperedge*. Given a collection $\mathcal{E} := (E_i)$ of m hyperedges, we refer to the pair $(\mathcal{V}, \mathcal{E})$ as a *hypergraph*. This hypergraph may be identified with an $m \times n$ *incidence matrix* of $\{0, 1\}$ entires, having no zero rows, as follows: the (i, j) entry is 1 if and only if $v_j \in E_i$, in which case we say row i is *incident* to column j, and that hyperedge E_i is incident to vertex v_j, and refer to (E_i, v_j) as an *incidence* of the hypergraph.

The number of hyperedges incident to a vertex is its *degree*. Fix a hypergraph $(\mathcal{V}, \mathcal{E})$. For $\mathcal{F} \subseteq \mathcal{E}$, the set $V(\mathcal{F}) \subseteq \mathcal{V}$ of vertices incident to at least one of the hyperedges in $\mathcal{F}$ is the *vertex span* of $\mathcal{F}$. We identify the hypergraph $(V(\mathcal{F}), \mathcal{F})$ by the edge subset $\mathcal{F}$ that induces it, and call $\mathcal{F} \subseteq \mathcal{E}$ a *partial hypergraph*. A partial hypergraph $\mathcal{F} \neq \emptyset$ is a *hypercycle* if every vertex v has even degree with respect to $\mathcal{F}$; this corresponds to a non-trivial left null vector for the incidence matrix of the hypergraph $(\mathcal{V}, \mathcal{E})$.

Given a hypergraph $(\mathcal{V}, \mathcal{E})$, the *2-core* is defined via the following algorithm:

1. If there exists no vertex of degree one, stop.
2. Otherwise, select an arbitrary vertex of degree one, and delete the unique incident hyperedge; then return to Step 1.

The algorithm terminates, because the partial hypergraphs are decreasing; the terminal partial hypergraph, which does not depend on the arbitrary choices made in Step 2 (see [8, pp. 127–128]), and may be empty, is called the *2-core* of $\mathcal{E}$, denoted $\mathrm{Core}(\mathcal{E})$.

The next result, Theorem 2.4, describes the 2-core of our random matrix $M(n, m_n)$: specifically, the limiting aspect ratio of the 2-core being less than or greater than 1

depends on the sign of $\psi(g^*(\alpha))$, where ψ and g^* are defined at (2.8) and (2.11) respectively. (A related result appears in [5].) While of interest in its own right, Theorem 2.4 has importance for the rank deficiency problem in view of the following observation: if the 2-core has more rows than columns, then the corresponding hypergraph has a hypercycle (see Lemma 5.1 below for details). Theorem 2.4 is thus the basis for the appearance in Theorem 2.2 of $\underline{\alpha}_\rho$ as defined at (2.12) (we explain in detail in Section 5).

We note, under the hypotheses of Theorem 2.4: (i) for $\alpha > \alpha_\rho^\sharp$, g^* is positive and strictly increasing; and (ii) g^* is right continuous, with a finite set of discontinuities $\mathcal{D}_\rho \subset (0,\infty)$ with $\alpha_\rho^\sharp = \min \mathcal{D}_\rho$. These and other facts are proved in Lemma 5.5 below.

Theorem 2.4. *Suppose W_n are uniformly bounded and $\mathbb{P}[W \geq 3] = 1$. Let $\alpha \in (0,\infty)$. Consider the 2-core of the random incidence matrix $M(n, m_n)$ where $m_n/n \to \alpha$ as $n \to \infty$. Then if $\alpha < \alpha_\rho^\sharp$, the number of rows in the 2-core is $o(n)$, a.s.*

Now suppose $\alpha > \alpha_\rho^\sharp$, so $g^* = g^*(\alpha) > 0$, and suppose that $\alpha \notin \mathcal{D}_\rho$. Then:

- *(i) n^{-1} times the number of rows in the 2-core converges a.s. to $\alpha\rho(g^*)$.*
- *(ii) n^{-1} times the number of occupied columns in the 2-core converges a.s. to $1 - \mathrm{e}^{-\nu}(1+\nu)$, where $\nu := \alpha\rho'(g^*)$.*
- *(iii) Almost surely, for all n large enough, the 2-core has more rows than occupied columns if $\psi(g^*(\alpha)) < 0$ but has fewer rows than occupied columns if $\psi(g^*(\alpha)) > 0$. Moreover, there exists $\delta > 0$ such that if $\alpha \in (\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$, for all n large enough, the 2-core has more rows than columns.*

In the example in Figure 1, and also in the fixed weight setting, $\psi(g^*(\alpha))$ changes sign only once, but in the general random weight setting it may change sign *multiple* times, leading, via Theorem 2.4, to non-monotonic behaviour for the aspect ratio of the 2-core. Figure 2 shows an example where $\rho(s) = 0.9183s^3 + 0.04s^{19} + 0.0417s^{41}$: as α increases from 0, the 2-core switches from having asymptotically more columns than rows to having more rows than columns not just once (at $\underline{\alpha}_\rho$), but *twice*, as $\psi(g^*(\alpha))$ changes sign. Proposition 5.7 below, and the subsequent discussion, explains some of the features in the figure. Thus the random weight setting displays subtle new phenomena not present in the fixed weight case that has been the focus of previous work.

2.4 Even occupancy in random allocations

One interpretation of the event $A(n,m)$ defined at (2.5) is in terms of the *random allocation model*. Suppose we have n urns, and for each row of M we allocate a collection of balls to a set of urns determined by the unit entries of that row of M. Event $A(n,m)$ is that all the urns end up with an even number of balls. Random allocations have been extensively studied; see e.g. [12, p. 101], and the monographs [16, 17, 20].

The following theorem, which we prove in Section 3, describes the exponential rate of decay for $\mathbb{P}_{\rho_n}[A(n,m_n)]$ where m_n/n has a finite positive limit. The theorem excludes the case in which *both* W and m_n only take odd values; if m is odd and W_n is odd a.s., then $\mathbb{P}_{\rho_n}[A(n,m)] = 0$ since the total number of units in the matrix is odd.

Theorem 2.5. *Suppose that $m_n/n \to \alpha \in (0,\infty)$ as $n \to \infty$, and that either (i) $m_n \in 2\mathbb{Z}$ for all n; or (ii) $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Then*

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}_{\rho_n}[A(n,m_n)] = -R_\rho(\alpha), \tag{2.14}$$

where $R_\rho(\alpha) > 0$ is continuous and nondecreasing in $\alpha > 0$ and is defined by

$$R_\rho(\alpha) := -\log \sup_{\gamma\in[0,1/2]} \left(\frac{(\rho(1-2\gamma))^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}} \right). \tag{2.15}$$

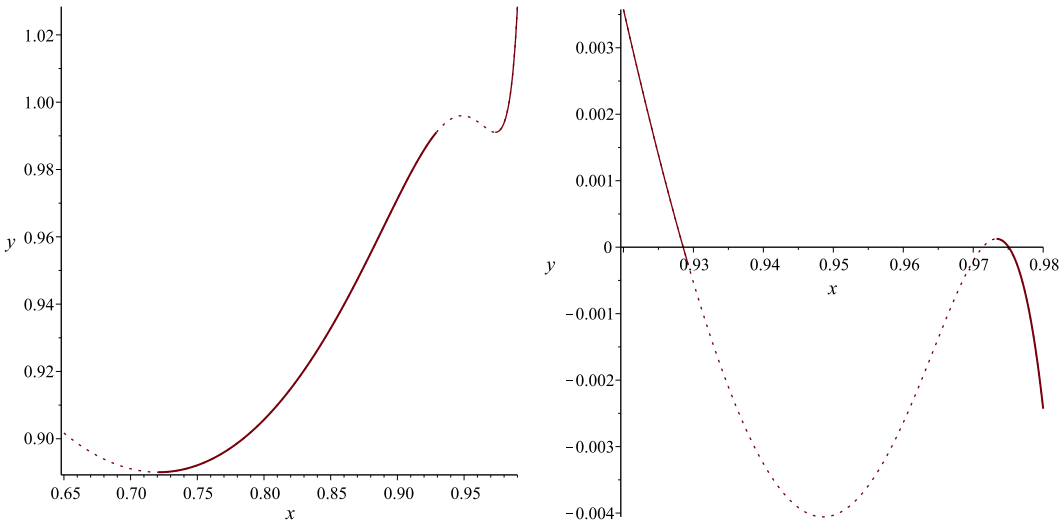

Figure 2: Example with $\rho(s) = 0.9183s^3 + 0.04s^{19} + 0.0417s^{41}$. The left plot shows parts of the curves $y = h(x)$ (all the line) and $x = g^*(y)$ (solid line). The right plot shows parts of the curves $y = \psi(x)$ (all the line) and the locus of $(g^*(\alpha), \psi(g^*(\alpha)))$ (solid line). Again $g^*(\alpha)$ has two discontinuities, one at $\alpha = \alpha_\rho^\sharp \approx 0.890061$ and one at $\alpha \approx 0.991044$, with the first corresponding to a jump from $g^* = 0$ to $g^* \approx 0.720793$ and the second to a jump from $g^* \approx 0.929269$ to $g^* \approx 0.973325$. The right plot shows the three positive zeros of $\psi(x)$. The two of these zeros achieved by $\psi(g^*(\alpha))$ are at $x \approx 0.928538$ and $x \approx 0.975069$. The first corresponds to $\alpha = \underline{\alpha}_\rho \approx 0.990686$ and the second to $\alpha \approx 0.991185$. Hence as α ranges in $(0, 1)$, $\psi(g^*(\alpha))$ changes sign from positive, to negative, to positive, and finally to negative again.

A consequence of Theorem 2.5 of independent interest concerns the probability $\pi_n(m)$ that all the components $Y_1, \dots, Y_n$ of a multinomial $(m; n^{-1}, \dots, n^{-1})$ random vector are even. Here Y_j can be interpreted as the occupancy of urn j after m balls are independently and uniformly distributed into n distinct urns: see e.g. [20, p. 11]. Then $\pi_n(m) = 2^{-n} \sum_{j=0}^{n} \binom{n}{j} (1 - (2j/n))^m$; this formula is known in the Ehrenfest urn literature [20, pp. 128–129] and can also be obtained from (3.1) below. If m is odd, $\pi_n(m)$ must be zero.

Proposition 2.6. *Let $\pi_n(m_n)$ denote the probability that all the n components of a multinomial $(m_n; n^{-1}, \dots, n^{-1})$ random vector are even. Suppose that m_n is even for each n and $m_n/n \to \alpha = \lambda \tanh \lambda \in (0, \infty)$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} n^{-1} \log \pi_n(m_n) = \log \cosh \lambda - (\lambda \tanh \lambda)(1 - \log \tanh \lambda). \tag{2.16}$$

We derive Proposition 2.6 from Theorem 2.5 in Section 3.5; it also follows from a result of Kolchin [15, Theorem 2, p. 141], and yet another proof is given in Section 3.6 of the first version of the present paper on ArXiv.

2.5 Thresholds in the fixed-weight case

In this section we consider the case where $\mathbb{P}[W = r] = 1$ for fixed $r \in \mathbb{N}$, which is the focus of the existing literature (see the discussion in Section 2.6 below). In particular, we discuss numerical and asymptotic evaluation of the thresholds α_r^*, $\alpha_r^\sharp$, and $\underline{\alpha}_r$, defined to be the values of α_ρ^*, $\alpha_\rho^\sharp$, and $\underline{\alpha}_\rho$, respectively, in the case where $\rho(s) = s^r$.

Appropriate versions of Theorems 2.1 and 2.2 apply in this setting. We remark that in the case $r = 2$, $\alpha_2^* = 1/2$ and the number of cycles $\mathcal{N}(n, m_n)$ in an Erdős–Rényi graph with $m_n/n \to \alpha$ has a Poisson limit with finite expectation for $\alpha \in (0, 1/2)$, but the limiting expectation is infinite for $\alpha \geq 1/2$ (see e.g. [16, §2.3]); we could not find in the literature an explicit reference to the fact that the expectation blows up *exponentially* with n for $\alpha > 1/2$, at the rate given by the appropriate case of Theorem 2.1.

Table 1 shows values of $\alpha_r^\sharp$, α_r^*, and $\underline{\alpha}_r$, for $r \leq 8$: we describe how these were computed in Appendix A, where we also review previous computations of these thresholds.

r	1	2	3	4	5	6	7	8
$\alpha_r^\sharp$	0	0.5	0.818469	0.772280	0.701780	0.637081	0.581775	0.534997
α_r^*	0	0.5	0.889493	0.967147	0.989162	0.996228	0.998650	0.999510
$\underline{\alpha}_r$	—	—	0.917935	0.976770	0.992438	0.997380	0.999064	0.999660

Table 1: Fixed row-weight thresholds. Note that $\underline{\alpha}_r$ is not defined when $r = 1$ or 2.

As suggested by the numerical results, it can be shown that, for r large enough, $\alpha_r^\sharp < \alpha_r^* < \underline{\alpha}_r < 1$; this is a consequence of the following result.

Proposition 2.7. *As $r \to \infty$,*

$$\alpha_r^\sharp \to 0; \quad 1 - \alpha_r^* \sim \frac{\mathrm{e}^{-r}}{\log 2}; \quad 1 - \underline{\alpha}_r \sim \mathrm{e}^{-r}. \tag{2.17}$$

The α_r^* result in (2.17) is due to Calkin [3]; we prove the other two in Appendix A.

2.6 Discussion and related results

2.6.1 Previous results on threshold values

In the simplest case, $W_n = W_n^{\mathrm{hyp}} := r \wedge n$ a.s., for a fixed $r \in \mathbb{N}$; then $W = r$ a.s. This fixed row weight 'hypergeometric' model is studied by Cooper [4]. A variation is the model in which r units are assigned to the row *independently* and uniformly at random, with multiplicities reduced mod 2. The latter 'binomial' model corresponds to $W_n = W_n^{\mathrm{bin}}$ distributed as the number of odd components in a multinomial $(r; n^{-1}, \ldots, n^{-1})$ random vector; then $W_n \xrightarrow{d} r$ (see Lemma 3.3 below). The $r \geq 3$ binomial model is studied by Kolchin [15]. Note that in this model rows of all zeroes may appear, in which case they are ignored (in other words, empty hyperedges are discounted): this is a small effect since $\mathbb{P}[W_n^{\mathrm{bin}} = 0] = O(n^{-r/2})$, so a vanishing proportion of rows needs to be discarded.

Phase transitions in the null vector problem for random matrices over finite fields with fixed row weight $r \geq 3$ have been studied since the early 1990s. In the case of the binomial model, the threshold α_r^*, $r \geq 3$, for $\mathbb{E}[\mathcal{N}(n, m_n)]$, $m_n/n \to \alpha$, was described by Balakin *et al.* [2] and Kolchin [15, 14]; in these results α_r^* is characterized by the fact that the expected number of non-trivial null vectors tends to 0 (∞) when $\alpha < \alpha_r^*$ ($\alpha > \alpha_r^*$), but the proofs show that the growth is in fact exponential for $\alpha > \alpha_r^*$. Calkin [3] and Cooper [4] also study α_r^*, $r \geq 3$, and in particular Calkin [3] studies α_r^* as $r \to \infty$; both [3] and [4] work in the case $W_n = r \wedge n$. Note that Cooper's [4] expression of the matrix problem is transposed compared to ours. Even the special case $\mathbb{P}[W = r] = 1$ of our Theorem 2.1 represents a slight generalization of the results just mentioned because it allows for any class of W_n provided $W_n \to r$ in probability.

In these previous investigations, the analytic description of the threshold α_r^* varies, but these descriptions can be shown to be consistent with ours: see Appendix A.

A similar tabulation to our tabulation of $\underline{\alpha}_r$ is given by Cooper [5, pp. 370–371], who also gives an equivalent analytic description of $\underline{\alpha}_r$ to our (A.2); see also Dietzfelbinger

et al. [10] which we discuss further in the next subsection. We note also that $\alpha_r^\sharp$ has received considerable attention in its own right: see e.g. [13] for its role in random XORSAT.

2.6.2 Between the two thresholds

The following problem arises in the XORSAT literature. Let $r \in \mathbb{N}$ with $r \geq 3$. Let M be our $m \times n$ matrix, with $m/n \to \alpha > 0$, and suppose $W_n = r$ a.s. for all $n \geq r$. Let N denote the number of column vectors $x \in \{0,1\}^n$ such that $Mx \equiv \omega$, where $\omega \in \{0,1\}^m$ is chosen uniformly at random (independent of M). Thus N is a random variable.

Dubois and Mandler [11] show for $r = 3$, and Dietzfelbinger *et al.* [10] extend to general $r \in \mathbb{N}$ with $r \geq 3$ (also providing a more detailed proof) the following result (see [11, Theorem 3.1] and [10, Theorem 1], and also [13]): there is a constant $\tilde{\alpha}_r > 0$ such that provided $\alpha < \tilde{\alpha}_r$, $\mathbb{P}[N > 0] \to 1$ as $n \to \infty$.

The proof of this in [11] is based on a second moment calculation. The analytical definition of $\tilde{\alpha}_r$ in [11, 10] is not obviously the same as our definition of $\underline{\alpha}_r$, but the definition in terms of cores (see [10, Proposition 3 and equation (4)]) seems to match our definition of $\underline{\alpha}_r$, and the numerical values in [10] are consistent with our $\underline{\alpha}_r$.

If we accept that $\tilde{\alpha}_r = \underline{\alpha}_r$, this result implies that if $\alpha < \underline{\alpha}_r$ there is, for n large enough, no non-zero left null vector for M, as follows. Suppose that a non-zero y satisfies $y \cdot M = \mathbf{0}$. Then $N > 0$ implies $y \cdot \omega = 0$. So $\mathbb{P}[y \cdot \omega = 0] \geq \mathbb{P}[N > 0] \to 1$, which contradicts the easy observation that $\mathbb{P}[y \cdot \omega = 0] = 1/2$ for non-zero y.

We may then deduce that in the case with $W_n = n \wedge r$, our Theorem 2.2 may be strengthened to $n^{-1}T_n \to \underline{\alpha}_r$ in probability. This implies that for α in the interval $(\alpha_r^*, \underline{\alpha}_r)$, a form of substantialism occurs; existence of any left null vector is unlikely, but if there is one, there are lots of them.

3 Multinomial parities and random allocations

3.1 Overview and terminology

In this section we work towards proving Theorem 2.5, in the context of the classical occupancy problems of random allocations of balls into urns.

We shall use the following terminology. Suppose W is a random variable taking values in $\mathbb{Z}_+$, and $k \in \mathbb{N}$, and $p, p_1, p_2, \ldots, p_k$ are numbers in $[0,1]$ such that $\sum_{i=1}^k p_i = 1$. (In most of the rest of the paper we assume $W \geq 1$, but for this section we can allow W to take value 0.) Let us say the random variable X has the $\mathrm{Bin}(W,p)$ distribution if for each $n \in \mathbb{Z}_+$ the conditional distribution of X, given that $W = n$, is binomial with parameters (n,p). Let us say that a random vector $(Z_1, \ldots, Z_k)$ has the multinomial $(W; p_1, \ldots, p_k)$ distribution if for each $n \in \mathbb{Z}_+$ the conditional distribution of $(Z_1, \ldots, Z_k)$, given that $W = n$, is multinomial with parameters $(n; p_1, \ldots, p_k)$.

As in Section 2.1, we assume W_n (having the distribution of row weights for our matrix with n columns) is chosen to converge in distribution to a limiting random variable W. An important special case is the so-called *binomial* model. In the binomial scheme take $W_n = W_n^{\mathrm{bin}}$ to be distributed as the number of odd components in a multinomial $(W; n^{-1}, \ldots, n^{-1})$ random vector. Note that we may also generate the corresponding row by first sampling the given multinomial vector and then reducing its elements mod 2. By Lemma 3.3 below, $W_n^{\mathrm{bin}} \xrightarrow{d} W$ as $n \to \infty$, so this is indeed a special case.

We write $\mathbb{P}_{\rho_n}^{\mathrm{bin}}$ for probability associated with the binomial allocation scheme. For the general model we write $\mathbb{P}_{\rho_n}$ as before.

3.2 Exact formulae for the allocation problem

Fix n. Let X_{ij} denote the jth component of X_i. Define the column sums Y_j and partial row sums $S_{i,J}$ of the matrix (X_{ij}) as follows (with standard addition):

$$Y_j := \sum_{i=1}^{m} X_{ij},\ j \in [n]; \quad \text{and} \quad S_{i,J} := \sum_{j \in J} X_{ij},\ J \subseteq [n].$$

Recall from (2.5) that $A(n,m)$ denotes the event that $\mathbf{1}$ is a null (row) vector for M.

Lemma 3.1. *In the binomial allocation scheme, we have the exact formula*

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m)] = 2^{-n} \sum_{j=0}^{n} \binom{n}{j} \left(\rho(1-(2j/n))\right)^m. \tag{3.1}$$

In the general allocation scheme,

$$\mathbb{P}_{\rho_n}[A(n,m)] = 2^{-n} \sum_{J \subseteq [n]} \left(\mathbb{E}_{\rho_n}[(-1)^{S_{1,J}}]\right)^m \tag{3.2}$$

$$= 2^{-n} \sum_{j=0}^{n} \binom{n}{j} (2p_j^{(n)} - 1)^m, \tag{3.3}$$

where $p_j^{(n)} := \sum_{r=0}^{n} p_{j,r} \mathbb{P}[W_n = r]$ *and* $p_{j,r}$ *is given by*

$$p_{j,r} = \frac{1}{\binom{n}{j}} \left(\binom{n-r}{j} + \binom{r}{2}\binom{n-r}{j-2} + \binom{r}{4}\binom{n-r}{j-4} + \cdots \right). \tag{3.4}$$

Proof. Event $A(n,m)$ occurs if and only if all the Y_j are even, so

$$\mathbb{P}[A(n,m)] = \mathbb{E} \prod_{j=1}^{n} \left(\frac{1+(-1)^{Y_j}}{2} \right) = 2^{-n} \sum_{J \subseteq [n]} \mathbb{E}\left[(-1)^{\sum_{j \in J} Y_j} \right],$$

where the latter sum is over subsets J of $[n]$, including the empty set. Since $\sum_{j \in J} Y_j = \sum_{i=1}^{m} S_{i,J}$ and $S_{1,J}, S_{2,J}, \ldots$ are i.i.d., (3.2) follows.

Consider the binomial allocation scheme. In the binomial model,

$$S_{1,J} = \sum_{j \in J} X_{1j} \equiv \sum_{j \in J} Z_j \pmod 2,$$

where $(Z_1, \ldots, Z_n)$ has a multinomial $(W; n^{-1}, \ldots, n^{-1})$ distribution so that $\sum_{j \in J} Z_j$ has a $\mathrm{Bin}(W, |J|/n)$ distribution. Recalling that if $\xi \sim \mathrm{Bin}(n,p)$ then $\mathbb{E}[s^\xi] = (sp + (1-p))^n$, we then obtain (3.1) from (3.2).

In the general scheme, conditional on $\sum_{j=1}^{n} X_{1j} = r$, the distribution of $S_{1,J}$ is hypergeometric with parameters $(n; |J|, r)$. Let $H \subseteq [n]$ denote the set of values of j for which $X_{1j} = 1$. For $r \in [n]$, we write $\mathbb{E}_r$ for expectation in the case where $\mathbb{P}[W_n = r] = 1$. Instead of fixing $J \subseteq [n]$ and choosing H as a uniform random r-subset, we obtain an exact formula for $\mathbb{E}_r[(-1)^{S_{1,J}}]$ by fixing $j = |J|$ and an r-subset H, and selecting J uniformly from the j-subsets of $[n]$. The probability $p_{j,r}$ that $S_{1,J} := |H \cap J|$ is even is given by summing probabilities for $|H \cap J| \in \{0, 2, 4, \ldots\}$, giving the expression in (3.4). It follows that $\mathbb{E}_r[(-1)^{S_{1,J}}] = 2p_{j,r} - 1$, and hence

$$\mathbb{E}_{\rho_n}[(-1)^{S_{1,J}}] = \sum_{r=1}^{n} (2p_{j,r} - 1)\mathbb{P}[W_n = r] = 2p_j^{(n)} - 1.$$

Substitution of this into (3.2) gives (3.3). □

3.3 Asymptotics in the binomial model

The remainder of Section 3 concerns asymptotic analysis of the quantities in Lemma 3.1. The first result enables us to work primarily with even m, which has technical advantages.

Lemma 3.2. *Suppose that $W_n \xrightarrow{d} W$ and $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Then for any $m > 3$,*

$$\log \mathbb{P}_{\rho_n}[A(n, m-3)] + O(\log n) \le \log \mathbb{P}_{\rho_n}[A(n,m)] \le \log \mathbb{P}_{\rho_n}[A(n, m+3)] + O(\log n).$$

Proof. The hypotheses imply that there exist $\varepsilon > 0$ and $r \in 2\mathbb{Z}$ such that $\mathbb{P}[W_n = r] > \varepsilon$ for all n large enough. For $m > 3$, suppose $A(n, m-3)$ occurs. Then $A(n,m)$ will occur if the 3 additional rows constitute a hypercycle. With probability at least ε^3, these new rows each have r units, and given this, there is a probability at least n^{-2r}, say, that these units form a hypercycle. Thus $\log \mathbb{P}_{\rho_n}[A(n,m)] \ge \log \mathbb{P}_{\rho_n}[A(n,m-3)] + O(\log n)$. Applying this inequality twice, once with $m+3$ in place of m, gives the result. □

Recall that in general we assume $W_n \xrightarrow{d} W$. Next we give an elementary lemma that confirms the binomial model's place in this framework.

Lemma 3.3. *For W a $\mathbb{Z}_+$-valued random variable, let W_n^{bin} be the number of odd components in a multinomial $(W; n^{-1}, \dots, n^{-1})$ random vector. Then $W_n^{\mathrm{bin}} \xrightarrow{d} W$ as $n \to \infty$.*

Proof. Let W and W_n^{bin} be coupled in the natural way. Then for each $k \in \mathbb{N}$, by the union bound $\mathbb{P}[W_n^{\mathrm{bin}} \ne W \mid W = k] \le n^{-1}\binom{k}{2}$, and the result follows easily from this and a truncation argument. □

We will prove Theorem 2.5 (in Section 3.5) by first showing that (2.14) holds in the binomial setting, using (3.1) and the Stirling approximation as discussed in Appendix B. Then we will extend this to the general setting using an approximation argument described in Section 3.4. We start with a slightly more general statement than (2.14) in the binomial case, which we will also need later in the proof of Theorem 2.1.

Lemma 3.4. *As defined at* (2.15)*, $R_\rho(\,\cdot\,)$ is continuous and nondecreasing. Suppose that either (i) $m_n \in 2\mathbb{Z}$ for all n; or (ii) $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Suppose that there exist α_1, α_2 with $0 < \alpha_1 < \alpha_2 < \infty$ such that, for all n sufficiently large, $\alpha_1 < m_n/n < \alpha_2$. Then,*

$$\begin{aligned} \limsup_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] &\le -R_\rho(\alpha_1); \\ \liminf_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] &\ge -R_\rho(\alpha_2). \end{aligned} \tag{3.5}$$

In particular, if $m_n/n \to \alpha > 0$, then

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] = -R_\rho(\alpha). \tag{3.6}$$

Proof. Suppose $m_n/n \in (\alpha_1, \alpha_2)$. First assume $m_n \in 2\mathbb{Z}$. By (3.1) and Lemma B.1(iii),

$$\begin{aligned} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m)] &\le (n+1)2^{-n} \max_{0\le j\le n} \binom{n}{j} |\rho(1-(2j/n))|^m \\ &\le (n+1)2^{-n} \sup_{\gamma\in[0,1/2]} \binom{n}{\gamma n} (\rho(1-2\gamma))^m, \end{aligned}$$

where we set $\binom{n}{x} = 0$ if x is a not an integer in $\{0, 1, \dots, n\}$. Using the upper bound on binomial coefficients from the first inequality in (B.1), we get

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \le (n+1) \sup_{\gamma\in[0,1/2]} \left(2\gamma^\gamma(1-\gamma)^{1-\gamma}\right)^{-n} (\rho(1-2\gamma))^{m_n}.$$

By monotonicity, we then obtain

$$n^{-1}\log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \le n^{-1}\log(n+1) + \log \sup_{\gamma\in[0,1/2]} g_{m_n/n}(\gamma), \tag{3.7}$$

where we have set

$$g_\alpha(\gamma) := \frac{(\rho(1-2\gamma))^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}};$$

so with R_ρ as defined at (2.15), $R_\rho(\alpha) = -\log\sup_{\gamma\in[0,1/2]} g_\alpha(\gamma)$. Note that, for $\alpha \ge 0$, $g_\alpha(\gamma)$ is continuous in $\gamma \in [0,1/2]$, and $g_\alpha(\gamma)$ is nonincreasing in α; this implies, by Dini's theorem, that if $\alpha' \to \alpha$ monotonically then $g_{\alpha'}$ converges uniformly to g_α on the compact interval $[0,1/2]$. It follows that $\alpha \mapsto \sup_{\gamma\in[0,1/2]} g_\alpha(\gamma)$ is continuous and nonincreasing. In particular, $R_\rho(\,\cdot\,)$ is continuous and nondecreasing, as claimed in the lemma. Moreover, we obtain from (3.7) and the fact that $m_n/n > \alpha_1$ that

$$\limsup_{n\to\infty} n^{-1}\log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \le \log \sup_{\gamma\in[0,1/2]} g_{\alpha_1}(\gamma),$$

which gives the first inequality in (3.5).

For the second inequality, we have from (3.1) and (B.2) that for any integer $i_n \le n/2$,

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \ge 2^{-n}\binom{n}{i_n}(\rho(1-(2i_n/n)))^{m_n} \ge \mathrm{e}^{-1/6}\left(\frac{n}{2\pi i_n(n-i_n)}\right)^{1/2}(g_{m_n/n}(i_n/n))^n,$$

using the fact that m_n is even. Then, since $m_n/n < \alpha_2$,

$$\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \ge \mathrm{e}^{-1/6}\left(\frac{n}{2\pi i_n(n-i_n)}\right)^{1/2}(g_{\alpha_2}(i_n/n))^n.$$

Now use continuity of g_α to choose a sequence of integers $i_n \le n/2$, $n \in \mathbb{N}$, such that $g_{\alpha_2}(i_n/n) \to \sup_{\gamma\in[0,1/2]} g_{\alpha_2}(\gamma)$, with $i_n \to \infty$ and $n - i_n \to \infty$ as $n \to \infty$. The lower bound in (3.5) follows, for m_n even.

The results in (3.5) extend to the case of odd m_n with $\mathbb{P}[W \in 2\mathbb{Z}] > 0$ by Lemma 3.2, which is applicable here by Lemma 3.3. Finally, (3.6) follows from (3.5) on taking $\alpha_1 = \alpha - \varepsilon$ and $\alpha_2 = \alpha + \varepsilon$, for arbitrary $\varepsilon > 0$, and using the continuity of R_ρ. □

3.4 Approximation by the binomial model

The exact formula (3.1) is simpler to work with than the more complicated exact formula (3.3), but intuition suggests that the asymptotics of any of the models in the class with $W_n \xrightarrow{d} W$ should be similar. The next result quantifies this intuition.

Lemma 3.5. *Suppose that $W_n \xrightarrow{d} W$ and either (i) $m_n \in 2\mathbb{Z}$ for all n; or (ii) $\mathbb{P}[W \in 2\mathbb{Z}] > 0$. Suppose that there exist α_1, α_2 with $0 < \alpha_1 < \alpha_2 < \infty$ and $n_0 \in \mathbb{N}$ such that $\alpha_1 < m_n/n < \alpha_2$ for all $n \ge n_0$. Then, uniformly over all such sequences m_n,*

$$\lim_{n\to\infty} n^{-1}\left|\log \mathbb{P}_{\rho_n}[A(n,m_n)] - \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\right| = 0. \tag{3.8}$$

In particular, if $m_n/n \to \alpha > 0$,

$$\lim_{n\to\infty} n^{-1}\log \mathbb{P}_{\rho_n}[A(n,m_n)] = \lim_{n\to\infty} n^{-1}\log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)].$$

Proof. Denote the weight of row 1 by W_n in the general allocation scheme, and by W_n^{bin} in the binomial scheme. Then $W_n \xrightarrow{d} W$ and $W_n^{\mathrm{bin}} \xrightarrow{d} W$, and we can work in a

probability space where $\mathbb{P}[W_n \neq W_n^{\mathrm{bin}}] \to 0$. As in Section 3.2, set $S_{1,J} = \sum_{j\in J} X_{1j}$; on $\{W_n = W_n^{\mathrm{bin}}\}$, the (conditional) law of $S_{1,J}$ is the same as in the binomial model. Thus

$$\sup_{J\subseteq[n]} \left|\mathbb{E}_{\rho_n}[(-1)^{S_{1,J}}] - \mathbb{E}_{\rho_n}^{\mathrm{bin}}[(-1)^{S_{1,J}}]\right| \leq 2\mathbb{P}[W_n \neq W_n^{\mathrm{bin}}] \to 0. \tag{3.9}$$

First assume $m_n \in 2\mathbb{Z}$. By (3.2) with (3.1) and (3.9), there exists a triangular array of numbers $(\delta_{j,n}, j \in [n]\cup\{0\}, n \in \mathbb{N})$ satisfying $\max_{0\leq j\leq n} |\delta_{j,n}| \to 0$ as $n \to \infty$, and

$$\mathbb{P}_{\rho_n}[A(n,m_n)] = 2^{-n}\sum_{j=0}^{n}\binom{n}{j}\left(\rho(1-(2j/n)) + \delta_{j,n}\right)^{m_n}. \tag{3.10}$$

Let $\varepsilon > 0$ and choose $K > 1$ large enough so that $\log(1-K^{-1}) > -\varepsilon/\alpha_2$ and $\log(1+K^{-1}) < \varepsilon/\alpha_2$. Then choose $\delta > 0$ such that $(K+1)\delta < \exp\{-1/(\alpha_1\varepsilon)\}$. Finally assume n is large enough so that $\sup_{j\in[n]\cup\{0\}} |\delta_{j,n}| \leq \delta$ and $\alpha_1 < (m_n/n) < \alpha_2$.

We split the sum in (3.10) into two parts, depending on the size of $\rho(1-(2j/n))$. First suppose that $|\rho(1-(2j/n))| \leq K\delta$. In this case

$$|(\rho(1-(2j/n)) + \delta_{j,n})^{m_n}| \leq ((K+1)\delta)^{m_n} \leq \exp\{-m_n/(\alpha_1\varepsilon)\} \leq \exp\{-n/\varepsilon\}, \tag{3.11}$$

and similarly,

$$(\rho(1-(2j/n)))^{m_n} \leq \exp\{-n/\varepsilon\}. \tag{3.12}$$

It follows from (3.10) and (3.11) that

$$\begin{aligned}\mathbb{P}_{\rho_n}[A(n,m_n)] = 2^{-n} &\sum_{j:|\rho(1-(2j/n))|>K\delta}\binom{n}{j}\left(\rho(1-(2j/n)) + \delta_{j,n}\right)^{m_n} \\ &+ O\left(\exp\{-n/\varepsilon\}\right).\end{aligned} \tag{3.13}$$

Now suppose that $|\rho(1-(2j/n))| > K\delta$. In this case

$$(\rho(1-(2j/n)) + \delta_{j,n})^{m_n} = (\rho(1-(2j/n)))^{m_n}\left(1+\theta_{j,n}K^{-1}\right)^{m_n},$$

where $|\theta_{j,n}| \leq 1$. By the choice of K, $\mathrm{e}^{-\varepsilon/\alpha_2} < 1+\theta_{j,n}K^{-1} < \mathrm{e}^{\varepsilon/\alpha_2}$, and hence

$$\exp\{-\varepsilon n\} < (1+\theta_{j,n}K^{-1})^{m_n} < \exp\{\varepsilon n\}.$$

Therefore for each j with $|\rho(1-(2j/n))| > K\delta$, there is $\varepsilon_{j,n}$ for which $|\varepsilon_{j,n}| < \varepsilon$, and

$$(\rho(1-(2j/n)) + \delta_{j,n})^{m_n} = (\rho(1-(2j/n)))^{m_n}\exp\{\varepsilon_{j,n}n\}.$$

Hence for the sum on the right-hand side of (3.13), there exists ε_n with $|\varepsilon_n| < \varepsilon$ so that

$$\begin{aligned}&2^{-n}\sum_{j:|\rho(1-(2j/n))|>K\delta}\binom{n}{j}\left(\rho(1-(2j/n)) + \delta_{j,n}\right)^{m_n} \\ &= 2^{-n}\exp\{\varepsilon_n n\}\sum_{j:|\rho(1-(2j/n))|>K\delta}\binom{n}{j}\left(\rho(1-(2j/n))\right)^{m_n},\end{aligned}$$

using the assumption that m_n is even so all the terms in the sum are nonnegative. Then by (3.12) and a similar argument to (3.13), the last displayed quantity is equal to $\mathbb{P}_{\rho_n}^{\mathrm{bin}}[A(n,m_n)]\exp\{\varepsilon_n n\} + O\left(\exp\{(\varepsilon-\varepsilon^{-1})n\}\right)$. Combining this with (3.13) we obtain

$$\mathbb{P}_{\rho_n}[A(n,m_n)] = \mathbb{P}_{\rho_n}^{\mathrm{bin}}[A(n,m_n)]\exp\{\varepsilon_n n\} + O\left(\exp\{(\varepsilon-\varepsilon^{-1})n\}\right), \tag{3.14}$$

uniformly in n (and m_n), the implicit constants depending on α_1 and α_2.

Writing $\Delta_n = O\left(\exp\{(\varepsilon - \varepsilon^{-1})n\}\right)$ for the final term in (3.14), we obtain

$$\log \mathbb{P}_{\rho_n}[A(n,m_n)] = \log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] + \varepsilon_n n + \log\left(1 + \frac{\Delta_n}{\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)]\exp\{\varepsilon_n n\}}\right).$$

By Lemma 3.4, we have that $\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] \geq \exp\{-nR_\rho(\alpha_2) - \varepsilon n\}$, for all n large enough. So we may take $\varepsilon > 0$ small enough so that the final $\log$ term in the last display is $O(\exp\{-n\})$, say. Hence

$$n^{-1}\log \mathbb{P}_{\rho_n}[A(n,m_n)] = n^{-1}\log \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,m_n)] + \varepsilon_n + o(1).$$

Since $|\varepsilon_n| \leq \varepsilon$ and $\varepsilon > 0$ was arbitrary, (3.8) follows in the case of even m_n. In the other case, Lemma 3.2 yields the same conclusion. The final statement in the lemma then follows from the final statement in Lemma 3.4. □

3.5 Proofs of Theorem 2.5 and Proposition 2.6

Now we can complete the proofs of Theorem 2.5 and Proposition 2.6.

Proof of Theorem 2.5. The theorem is now a consequence of Lemmas 3.4 and 3.5. □

Proof of Proposition 2.6. Set

$$f_\alpha(\gamma) := \log\left(\frac{(1-2\gamma)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right). \tag{3.15}$$

In the case $\rho(s) = s$, it follows from Theorem 2.5 that $n^{-1}\log \pi_n(m_n) \to \sup_{\gamma\in[0,1/2]} f_\alpha(\gamma)$. Proposition 2.6 will follow once we prove that, setting $\alpha = \lambda\tanh\lambda$,

$$\sup_{\gamma\in[0,1/2]} f_\alpha(\gamma) = -(\lambda\tanh\lambda)(1 - \log(\tanh\lambda)) + \log(\cosh\lambda). \tag{3.16}$$

Note that $f_\alpha(\gamma) \to -\infty$ as $\gamma \uparrow 1/2$. Differentiating (3.15) gives, for $\gamma \in (0,1/2)$,

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} f_\alpha(\gamma) = -\frac{2\alpha}{1-2\gamma} + \log\left(\frac{1-\gamma}{\gamma}\right),$$

which is zero at $\gamma_1 := \gamma_1(\alpha) \in (0,1/2)$ defined implicitly in terms of α by

$$\alpha = \frac{1}{2}(1-2\gamma_1)\log\left(\frac{1-\gamma_1}{\gamma_1}\right). \tag{3.17}$$

For $\alpha > 0$ (3.17) defines a unique stationary value $\gamma_1 \in (0,1/2)$ which is a local maximum, since for $\gamma_1 \in (0,1/2)$ the right-hand side of (3.17) is positive, continuous, and strictly decreasing as a function of γ_1, vanishing at $\gamma_1 = 1/2$; this local maximum is indeed the maximum of $f_\alpha(\gamma)$ for $\gamma \in [0,1/2]$ since $f'_\alpha(\gamma) \to \infty$ as $\gamma \downarrow 0$ (and also $f''_\alpha(\gamma_1) < 0$).

Setting $\lambda = \frac{1}{2}\log\left(\frac{1-\gamma_1}{\gamma_1}\right)$ we see that $\lambda\tanh\lambda = \alpha$ as given by (3.17), since we get $\tanh\lambda = 1 - 2\gamma_1$. To verify (3.16) we need to express $f_\alpha(\gamma_1)$ in terms of λ to get the expression on the right-hand side of (3.16). We have

$$\begin{aligned} f_\alpha(\gamma_1) &= \alpha\log(1-2\gamma_1) - \log 2 - \gamma_1\log\gamma_1 - (1-\gamma_1)\log(1-\gamma_1) \\ &= (\lambda\tanh\lambda)\log\tanh\lambda + \log\cosh\lambda + (\tfrac{1}{2}-\gamma_1)\log\gamma_1 + (\gamma_1 - \tfrac{1}{2})\log(1-\gamma_1), \end{aligned}$$

using the fact that $\log\tanh\lambda = \log(1-2\gamma_1)$ and $\log\cosh\lambda = -\frac{1}{2}\log(1-\tanh^2\lambda) = -\log 2 - \frac{1}{2}\log\gamma_1 - \frac{1}{2}\log(1-\gamma_1)$. The terms involving γ_1 in the last displayed equation simplify to $-\lambda\tanh\lambda = -\alpha$ as given by (3.17). Thus we verify (3.16). □

4 The expected number of null vectors

4.1 Exact formula for the expected number of null vectors

Let $\mathcal{N}(n,m;\ell)$ denote the number of left null vectors of weight ℓ, so that

$$\mathcal{N}(n,m) = \sum_{\ell=0}^{m} \mathcal{N}(n,m;\ell). \tag{4.1}$$

By the argument around (2.6), we have that

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m;\ell)] = \binom{m}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)]. \tag{4.2}$$

The proof of Theorem 2.1 is based on an asymptotic analysis of (2.6) with our exact formulae for $\mathbb{P}_{\rho_n}[A(n,\ell)]$. As in the proof of Theorem 2.5 (see Section 2.4) it is most convenient to work in the binomial model, for which W_n^{bin} is the number of odd components in a multinomial $(W; n^{-1}, \dots, n^{-1})$ vector. Thus a key step in the proof will be showing that, in the general case of $W_n \xrightarrow{d} W$, the expression in (2.6) can be well approximated by the binomial case. First, in the next section, we make some preliminary computations.

4.2 Analytic preliminaries

Before embarking on the main proof, we study the rate functions that will appear. Define

$$F_{\rho,\alpha}(\gamma) := \log\left(\frac{(1+\rho(1-2\gamma))^{\alpha}}{2\gamma^{\gamma}(1-\gamma)^{1-\gamma}}\right), \tag{4.3}$$

and recall from (2.1) and (2.2) that $F_\rho(\alpha) = \sup_{\gamma\in[0,1/2]} F_{\rho,\alpha}(\gamma)$ and $\alpha_\rho^* = \inf\{\alpha \geq 0 : F_\rho(\alpha) > 0\}$. Note that for $\gamma \in [0,1/2]$, $\rho(1-2\gamma) \geq 0$. By continuity, $F_{\rho,\alpha}(\gamma)$ attains its supremum over $\gamma \in [0,1/2]$; we denote by $\gamma_0 := \gamma_0(\alpha) \in [0,1/2]$ the *smallest* point at which the supremum is attained. The next lemma collects results on $F_\rho(\alpha)$ and α_ρ^*.

Lemma 4.1. *Suppose that $\mathbb{P}[W=0] = 0$. For any $\alpha \geq 0$, $F_\rho(\alpha) \geq 0$, and F_ρ is continuous and nondecreasing. The threshold α_ρ^* enjoys the following properties.*

- *(i)* $\alpha_\rho^* \in [0,1]$, *and* $F_\rho(\alpha) = 0$ *for* $\alpha \leq \alpha_\rho^*$ *but* $F_\rho(\alpha) > 0$ *for* $\alpha > \alpha_\rho^*$.
- *(ii) If* $\alpha < \alpha_\rho^*$, *then* $\gamma_0(\alpha) = 1/2$ *and for any* $\varepsilon > 0$, $\sup_{\gamma\in[0,(1/2)-\varepsilon]} F_{\rho,\alpha}(\gamma) < 0$.
- *(iii) If* $\alpha > \alpha_\rho^*$, *then* $\gamma_0(\alpha) \in [0,1/2)$.
- *(iv) Suppose that $\tilde{W}$ is another $\mathbb{N}$-valued random variable, with $\tilde{\rho}(s) = \mathbb{E}[s^{\tilde{W}}]$, such that $\tilde{\rho}(s) \leq \rho(s)$ for all $s \in [0,1]$. Then $\alpha_{\tilde{\rho}}^* \geq \alpha_\rho^*$.*
- *(v)* $\alpha_\rho^* = 0$ *if and only if* $\mathbb{P}[W=1] > 0$.
- *(vi) If* $\mathbb{P}[W=2] = 1$, *then* $\alpha_\rho^* = 1/2$.
- *(vii) If* $\mathbb{E}[W] < \infty$, *then* $\alpha_\rho^* < 1$.

Proof. By Lemma B.1, $\rho(0) = \mathbb{P}[W=0] = 0$ and $\rho(1) = 1$; hence $F_{\rho,\alpha}(1/2) = 0$ and $F_{\rho,\alpha}(0) = (\alpha-1)\log 2$, so that $F_\rho(\alpha) \geq (\alpha-1)^+ \log 2 \geq 0$. Since $F_{\rho,\alpha}(\gamma)$ is nondecreasing as a function of $\alpha \geq 0$, Dini's theorem implies continuity of $F_\rho(\alpha)$ as a function of $\alpha \geq 0$.

For part (i), $F_\rho(\alpha) \geq (\alpha-1)^+ \log 2$ implies that $F_\rho(\alpha) > 0$ for $\alpha > 1$, so that $\alpha_\rho^* \leq 1$. On the other hand, for $\alpha < 1$, $\gamma_0(\alpha) \in (0,1/2]$, since by continuity there is some neighbourhood of 0 for which $F_{\rho,\alpha}(\gamma) < 0$.

Since $F_{\rho,\alpha}(\gamma)$ is nondecreasing as a function of α, for $\alpha' \geq \alpha$, $F_\rho(\alpha') \geq F_{\rho,\alpha'}(\gamma_0(\alpha)) \geq F_\rho(\alpha)$, i.e., F_ρ is also nondecreasing. Hence $F_\rho(\alpha) > 0$ for $\alpha > \alpha_\rho^*$. Also, the fact that $F_\rho(\alpha) = 0$ for $\alpha < \alpha_\rho^*$ is immediate from the definition of α_ρ^* and the fact that $F_\rho(\alpha) \geq 0$. Then $F_\rho(\alpha_\rho^*) = 0$ by the continuity of F_ρ established above. Thus we obtain part (i).

For part (ii), suppose that $\alpha < \alpha^*_\rho$. Suppose for some $\gamma_0 \in [0, 1/2)$ that $F_{\rho,\alpha}(\gamma_0) \geq 0$. Since $\rho(1-2\gamma) > 0$ for $\gamma < 1/2$, $F_{\rho,\alpha}(\gamma_0)$ is strictly increasing in α, so there exists $\alpha' \in (\alpha, \alpha^*_\rho)$ for which $F_{\rho,\alpha'}(\gamma_0) > 0$, contradicting the definition of α^*_ρ. This gives (ii).

For part (iii), suppose that $\alpha > \alpha^*_\rho$. Then $F_\rho(\alpha) > 0$ by part (i) of the lemma; since $F_{\rho,\alpha}(1/2) = 0$, the supremum is attained in $[0, 1/2)$.

For part (iv), we have that for any $\gamma \in [0, 1/2]$, $F_{\rho,\alpha}(\gamma) \geq F_{\tilde{\rho},\alpha}(\gamma)$, since $\rho(1-2\gamma) \geq \tilde{\rho}(1-2\gamma)$. So $F_\rho(\alpha) \geq F_{\tilde{\rho}}(\alpha)$ for all $\alpha \geq 0$, and hence $\alpha^*_{\tilde{\rho}} \geq \alpha^*_\rho$.

For the remaining parts of the lemma we use more detailed properties of the generating function $\rho(s)$ (see Lemma B.1). For part (v), differentiating in (4.3) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} F_{\rho,\alpha}(\gamma) = -\frac{2\alpha\rho'(1-2\gamma)}{1+\rho(1-2\gamma)} + \log\left(\frac{1-\gamma}{\gamma}\right); \tag{4.4}$$

this is well defined at least for $\gamma \in (0,1)$. At $\gamma = 1/2$ this equates to $-2\alpha\mathbb{P}[W=1]$, since by Lemma B.1 $\rho'(0) = \mathbb{P}[W=1]$ and $\rho(0) = 0$. So if $\mathbb{P}[W=1] > 0$, $F_{\rho,\alpha}(\gamma)$ is equal to 0 at $\gamma = 1/2$ and its derivative there is negative for any $\alpha > 0$, so that, for any $\alpha > 0$, $F_{\rho,\alpha}(\gamma) > 0$ for some $\gamma < 1/2$. The 'if' part of part (v) follows.

Conversely, suppose that $\mathbb{P}[W=1] = 0$. Then the previous argument shows that $F_{\rho,\alpha}(1/2) = F'_{\rho,\alpha}(1/2) = 0$, while a calculation shows that $F''_{\rho,\alpha}(1/2) = 4\alpha\rho''(0) - 4$. Hence by continuity there exists $\delta > 0$ such that for $\alpha < \delta$ and $(1/2) - \delta \leq \gamma \leq 1/2$ we have $F''_{\rho,\alpha}(\gamma) \leq -3$. Hence by Taylor's theorem, $F_{\rho,\alpha}(\gamma) \leq 0$ for $\alpha < \delta$ and $(1/2) - \delta \leq \gamma \leq 1/2$. Also, $F_{\rho,\alpha}(\gamma) \to -\log(2\gamma^\gamma(1-\gamma)^{1-\gamma})$ as $\alpha \to 0$, which is strictly negative apart from at $\gamma = 1/2$. Thus by Dini's theorem, for all α small enough we have $F_{\rho,\alpha}(\gamma) \leq 0$ for $\gamma \leq (1/2) - \delta$. So all together we have shown that $F_{\rho,\alpha}(\gamma) \leq 0$ for all α sufficiently small. Hence $\alpha^*_\rho > 0$ in this case, giving the 'only if' part of (v).

For part (vi), suppose that $\mathbb{P}[W=2] = 1$, i.e., $\rho(s) = s^2$. In this case, (4.4) has a zero at $\gamma \in [0, 1/2)$ if $\alpha = s(\gamma)$ where

$$s(\gamma) = \frac{1+(1-2\gamma)^2}{4(1-2\gamma)} \log\left(\frac{1-\gamma}{\gamma}\right).$$

We claim that $s(\gamma)$ is decreasing on $[0, 1/2)$, with a unique minimum of $s(1/2) = 1/2$. To verify this, we show $s'(\gamma) < 0$ for $\gamma \in [0, 1/2)$, which, after simplification, amounts to

$$\frac{(1+(1-2\gamma)^2)(1-2\gamma)}{8\gamma^2(1-\gamma)^2} > \log\left(\frac{1-\gamma}{\gamma}\right).$$

Setting $z = 1-2\gamma$, it suffices to show that $\frac{z}{(1-z^2)^2} > \frac{1}{2}\log\left(\frac{1+z}{1-z}\right)$ for $z \in (0,1]$, which can be verified by term-by-term comparison of the corresponding power series, namely $z + 2z^3 + 3z^5 + \cdots > z + \frac{z^3}{3} + \frac{z^5}{5} + \cdots$. Hence $s(\gamma) = \alpha$ has no solution for $\alpha < 1/2$, in which case the only stationary value of $F_{\rho,\alpha}$ is at $\gamma = 1/2$, necessarily the maximum. Hence $\alpha^*_\rho \geq 1/2$. On the other hand, if $\alpha > 1/2$ then $F''_{\rho,\alpha}(1/2) = 8\alpha - 4 > 0$, while $F_{\rho,\alpha}(1/2) = F'_{\rho,\alpha}(1/2) = 0$, so by Taylor's theorem and continuity there exists $\gamma < 1/2$ with $F_{\rho,\alpha}(\gamma) > 0$. Hence $\alpha^*_\rho = 1/2$, proving (vi).

Finally we prove part (vii). If $\mathbb{E}[W] < \infty$, Lemma B.1(ii) implies that, as $\gamma \downarrow 0$, $\rho'(1-2\gamma) = \mathbb{E}[W] + o(1)$. Thus the final term on the right-hand side of (4.4) dominates in the $\gamma \downarrow 0$ limit, and there exists $\delta > 0$ such that $\frac{\mathrm{d}}{\mathrm{d}\gamma} F_{\rho,\alpha}(\gamma) \geq \delta$ for all $\gamma \in [0, \delta]$ and all $\alpha \in [0,1]$. Then by an application of the mean value theorem, $F_{\rho,\alpha}(\delta) \geq (\alpha - 1)\log 2 + \delta^2$ for all $\alpha \in [0,1]$. Thus taking $\alpha < 1$ close enough to 1 we see that $F_{\rho,\alpha}(\delta) > 0$, which implies that $\alpha^*_\rho < 1$. □

4.3 Approximation by the binomial model

In Section 3.4 we showed (in Lemma 3.5) that $\mathbb{P}_{\rho_n}[A(n,m_n)]$ can be well approximated by $\mathbb{P}^{\rm bin}_{\rho_n}[A(n,m_n)]$ on the logarithmic scale, provided that $m_n/n \to \alpha$. The following result is an analogous approximation lemma for $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)]$. One could obtain such a result from Lemma 3.5 applied to (2.6), with some work (including dealing separately with terms with $\ell = o(n)$: cf Section 4.4 below). However, it is more convenient to proceed directly, albeit using similar ideas to the proof of Lemma 3.5; helpful is the fact that $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]$ possesses monotonicity properties absent for $\mathbb{P}_{\rho_n}[A(n,m)]$.

Lemma 4.2. *Suppose that $W_n \xrightarrow{d} W$ and $m_n/n \to \alpha > 0$. Then*

$$\lim_{n\to\infty} n^{-1} |\log \mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)] - \log \mathbb{E}^{\rm bin}_{\rho_n}[\mathcal{N}(n,m_n)]| = 0.$$

Proof. We use a coupling argument, constructing the general model with row weights distributed as $W_n \xrightarrow{d} W$ on the same probability space as the binomial model with row weights distributed as $W^{\rm bin}_n \xrightarrow{d} W$ (cf Lemma 3.3). We can use a probability space in which, for each row, the weight in each model converges almost surely to a copy of W. Indeed, let $W(1), W(2), \ldots$ be independent copies of W. Using the Skorokhod representation theorem, we may take $W_{n,1}, W_{n,2}, \ldots$ as independent copies of W_n, being the weights of the rows in the general model, such that $W_{n,i} \to W(i)$ almost surely. Also, take $W^{\rm bin}_{n,i}$ to be the number of odd components in a multinomial $(W(i); n^{-1}, \ldots, n^{-1})$ distribution, so that $W^{\rm bin}_{n,1}, W^{\rm bin}_{n,2}, \ldots$ are independent copies of $W^{\rm bin}_n$ and the weights of the rows in the binomial model. We also couple the row *entries*: if $W^{\rm bin}_{n,i} = W_{n,i}$, we generate a single row i with the given weight to use in both models, otherwise, it suffices to generate the two rows independently given their (different) weights.

Take $\varepsilon > 0$. Let $A_n(i) := \{W_{n,i} \neq W^{\rm bin}_{n,i}\}$. Then for any $\delta > 0$, we may take n large enough so that $\mathbb{P}[A_n(i)] \leq \delta$, uniformly in i. Let $K(n,m) = \sum_{i=1}^m \mathbf{1}_{A_n(i)}$ denote the number of 'bad' rows. Then $K(n,m)$ is stochastically dominated by a $\mathrm{Bin}(m,\delta)$ variable. In particular, for any fixed $\varepsilon > 0$ and any $C < \infty$, standard binomial tail bounds imply that we may take δ small enough, and hence n sufficiently large, so that

$$\mathbb{P}[K(n,m_n) \geq \varepsilon n] \leq \mathbb{P}[\mathrm{Bin}(2\alpha n, \delta) \geq \varepsilon n] \leq \exp\{-Cn\}. \tag{4.5}$$

We claim that each row added to a matrix can increase the number of null vectors by at most a factor of 2; this follows from (1.1) and (1.2). Hence

$$2^{-K(n,m_n)} \leq \frac{\mathcal{N}(n,m_n)}{\mathcal{N}'(n,m_n)} \leq 2^{K(n,m_n)}, \tag{4.6}$$

where $\mathcal{N}(n,m_n)$ is the number of null vectors in the matrix with the $W_{n,i}$ and $\mathcal{N}'(n,m_n)$ is the number of null vectors in the matrix with the $W^{\rm bin}_{n,i}$. Partitioning according to whether $K(n,m_n) \leq \varepsilon n$ or not, and using the crude bound $\mathcal{N}(n,m_n) \leq 2^{m_n}$ when not, we obtain from (4.6) the upper bound

$$\begin{aligned}\mathbb{E}[\mathcal{N}(n,m_n)] &\leq 2^{\varepsilon n}\mathbb{E}[\mathcal{N}'(n,m_n)] + 2^{m_n}\mathbb{P}[K(n,m_n) > \varepsilon n] \\ &\leq 2^{\varepsilon n}\mathbb{E}[\mathcal{N}'(n,m_n)] + 2^{\varepsilon n},\end{aligned}$$

say, for all n large enough, by (4.5). Since $\mathcal{N}'(n,m_n) \geq 1$, we may divide through by $\mathbb{E}[\mathcal{N}'(n,m_n)]$ and take logs in the last display to obtain

$$\limsup_{n\to\infty} n^{-1}\left(\log \mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n)] - \log \mathbb{E}^{\rm bin}_{\rho_n}[\mathcal{N}(n,m_n)]\right) = 0,$$

using the fact that $\varepsilon > 0$ was arbitrary. A similar argument using (4.6) in the other direction yields the complementary $\liminf$ result, giving the statement in the lemma. □

4.4 Null vectors consisting of few rows

In the asymptotics of $\mathbb{E}_{\rho_n}[\mathcal{N}(n,m)]$, it turns out that null vectors of *low weight* play a distinct and important role. Recall (4.1). The main result of this section is the following lemma, which exhibits a polynomial growth rate for null vectors of few rows.

Lemma 4.3. *Suppose that there exist $r_0 \geq 3$ and $r_1 < \infty$ such that $\mathbb{P}[r_0 \leq W_n \leq r_1] = 1$ for all n. Suppose that $m_n/n \to \alpha > 0$. Then there exists $\delta > 0$ such that*

$$\sum_{2 \leq \ell \leq \delta n} \mathbb{E}_{\rho_n}[\mathcal{N}(n, m_n; \ell)] = O(n^{2-r_0}). \tag{4.7}$$

Remark 4.4. *The exponent $2 - r_0$ in (4.7) cannot be improved when $\mathbb{P}[W_n = r_0] > 0$, because $\mathbb{E}[\mathcal{N}(n, m_n; 2)]$ is itself of order n^{2-r_0}. Indeed, there are of order m_n^2 weight-2 candidate vectors, and each is null if each of the two corresponding rows have r_0 non-zeros in matching positions, an event of probability of order n^{-r_0}.*

Proof of Lemma 4.3. Let $n, \ell \in \mathbb{N}$. Let $R = R(n,\ell)$ denote the 'column range' of $M(n,\ell)$, i.e., the number of columns of non-zero degree. We estimate $\mathbb{P}_{\rho_n}[A(n,\ell)]$ by considering separately the events $R \leq k$ and $R > k$, where $k = k(\ell) \in [n]$ is to be chosen later.

We describe $M(n,\ell)$ in the language of allocations: for each row, a W_n-distributed collection of balls is distributed uniformly at random among n urns (columns), at most one ball per urn. If $R \leq k$ then there is some set of k urns that contain all the balls. For each ball, the probability that it lands in one of the first k urns, given that the other balls cast so far for that row all land in the first k urns, is at most k/n. Hence since for each of the ℓ rows at least r_0 balls are cast,

$$\mathbb{P}_{\rho_n}[R \leq k] \leq \binom{n}{k}\left(\frac{k}{n}\right)^{\ell r_0} \leq \frac{n^{k-\ell r_0} k^{\ell r_0}}{k!}. \tag{4.8}$$

For $A(n,\ell)$ to occur, each of the columns in the range must have degree at least 2. Thus if $R > k$ and $A(n,\ell)$ occurs there is a collection of $k+1$ urns such that each urn in the collection gets at least 2 balls. Let $B(i)$ be the event that urn i gets at least 2 balls. The probability that a particular entry is 1, given the values of up to k other entries in the same row, is at most $r_1/(n-k)$. Hence the union bound yields for $1 \leq j \leq k+1$ that

$$\mathbb{P}_{\rho_n}[B(j) \mid \cap_{i=1}^{j-1} B(i)] \leq \binom{\ell}{2}\left(\frac{r_1}{n-k}\right)^2,$$

and hence we have

$$\mathbb{P}_{\rho_n}[\cap_{i=1}^{k+1} B(i)] \leq \left(\binom{\ell}{2}\left(\frac{r_1}{n-k}\right)^2\right)^{k+1}.$$

Hence by the union bound, provided $k \leq n/2$ we have

$$\mathbb{P}_{\rho_n}[\{R > k\} \cap A(n,\ell)] \leq \binom{n}{k+1}\left(\binom{\ell}{2}\left(\frac{2r_1}{n}\right)^2\right)^{k+1} \leq \frac{n^{-(k+1)}\ell^{2(k+1)}c_1^{k+1}}{(k+1)!},$$

where we put $c_1 = 2r_1^2$. Combined with (4.8) this gives

$$\mathbb{P}_{\rho_n}[A(n,\ell)] \leq \frac{n^{k-\ell r_0}k^{\ell r_0}}{k!} + \frac{n^{-(k+1)}\ell^{2(k+1)}c_1^{k+1}}{(k+1)!}.$$

For all n large enough, $m_n \le (1+\alpha)n$ so that, for all ℓ, and for $k \le n/2$, by (4.2),

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n;\ell)] = \binom{m_n}{\ell}\mathbb{P}_{\rho_n}[A(\ell,n)] \\ \le \left(\frac{((\alpha+1)n)^\ell}{\ell!}\right)\left(\frac{n^{k-\ell r_0}k^{\ell r_0}}{k!} + \frac{n^{-(k+1)}\ell^{2(k+1)}c_1^{k+1}}{(k+1)!}\right). \tag{4.9}$$

Taking $k = \ell + r_0 - 2$, we obtain for each fixed ℓ that for some constant $c(\ell)$ we have

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n;\ell)] \le c(\ell)(n^{(\ell-1)(2-r_0)} + n^{1-r_0}),$$

which is $O(n^{2-r_0})$ for any fixed $\ell \ge 2$.

Fix an integer $K \ge 2$, to be chosen later, and take $K \le \ell \le \delta n$. Now put $k = \ell(r_0-1) - \lceil \ell/2 \rceil$. Assume $\delta \le 1/(2r_0)$; then for $\ell \le \delta n$ this choice of k satisfies $k \le n/2$, so (4.9) remains valid. Also, since $r_0 \ge 3$, $k \ge \frac{3\ell}{2} - 1 \ge \ell$ provided $\ell \ge 2$.

By the bound $\mathrm{e}^\ell \ge \frac{\ell^\ell}{\ell!}$ and similar for k, there are constants c_2, c_3, c_4 such that the first term in the right side of (4.9) (i.e. the product of the first factor with the first term in the second factor) is bounded by a constant times

$$\frac{n^{\ell(1-r_0)+k}k^{\ell r_0}c_2^\ell}{\ell^\ell k^k} \le \frac{n^{-\lceil \ell/2 \rceil}\ell^{r_0\ell}c_3^\ell}{\ell^\ell \ell^{(r_0-1)\ell - \lceil \ell/2 \rceil}} = \left(\frac{c_4\ell}{n}\right)^{\lceil \ell/2 \rceil}. \tag{4.10}$$

Similarly, there are constants c_5, c_6, c_7 such that the second term in the right side of (4.9) is bounded by a constant times

$$\frac{n^{\ell-k-1}\ell^{2\ell(r_0-1)-2\lceil \ell/2 \rceil+2}c_5^\ell}{\ell^\ell(k+1)^{k+1}} \le \frac{n^{\ell(2-r_0)+\lceil \ell/2 \rceil}\ell^{\ell(2r_0-2)-2\lceil \ell/2 \rceil}c_6^\ell\ell^2}{\ell^{\ell r_0 - \lceil \ell/2 \rceil + 1}n} \\ \le \left(\frac{c_7\ell}{n}\right)^{\ell(r_0-2)-\lceil \ell/2 \rceil}(\ell/n).$$

Combining with (4.10), since $r_0 \ge 3$ so $r_0 - 2 \ge 1$ and $(\ell/2) \le \lceil \ell/2 \rceil \le (\ell/2)+1$, we can find a constant c_8 such that for $2 \le \ell \le n$ we have

$$\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n;\ell)] \le c_8\left(\frac{c_8\ell}{n}\right)^{\ell/2}.$$

By calculus we have that $\left(\frac{c_8 x}{n}\right)^x$ is decreasing in $x \le n/(c_8\mathrm{e})$, so provided $\delta \le 1/(c_8\mathrm{e})$ the last bound is maximized, over $K \le \ell \le \delta n$, at $\ell = K$, so that

$$\sum_{K \le \ell \le \delta n}\mathbb{E}_{\rho_n}[\mathcal{N}(n,m_n;\ell)] \le c_8\delta n\left(\frac{c_8K}{n}\right)^{K/2},$$

which is $O(n^{2-r_0})$ provided we choose K so that $K/2 \ge r_0 - 1$. □

4.5 Proof of Theorem 2.1

First we prove Theorem 2.1 for the binomial model, i.e., for $\mathbb{P}^{\mathrm{bin}}_{\rho_n}$ on the right-hand side of (2.6), and then use Lemma 4.2. Specifically, we prove the following result.

Lemma 4.5. *Suppose that $\mathbb{P}[W \ge 1] = 1$. Suppose that $m_n/n \to \alpha \in (0,\infty)$ as $n \to \infty$. Then with $F_\rho(\alpha)$ as defined by* (2.2)*,*

$$\lim_{n\to\infty} n^{-1}\log \mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] = F_\rho(\alpha). \tag{4.11}$$

Proof. From 3.1,

$$\sum_{\ell=0}^{m}\binom{m}{\ell}\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] \le (m+1)(n+1)2^{-n}\sup_{0\le\ell\le m}\sup_{0\le j\le n}\binom{m}{\ell}\binom{n}{j}|\rho(1-(2j/n))|^{\ell}$$
$$\le (m+1)(n+1)2^{-n}\sup_{\beta\in[0,1]}\sup_{\gamma\in[0,1]}\binom{m}{\beta m}\binom{n}{\gamma n}|\rho(1-2\gamma)|^{\beta m}, \qquad (4.12)$$

setting $\binom{n}{x}=0$ for $x\notin\{0,1,\dots,n\}$. Write

$$S_\alpha(\beta,\gamma) := \left(\frac{|\rho(1-2\gamma)|^\beta}{\beta^\beta(1-\beta)^{1-\beta}}\right)^\alpha\left(\frac{1}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right).$$

Taking $m=m_n=O(n)$ in (4.12) and using the first inequality in (B.1), we obtain

$$n^{-1}\log\sum_{\ell=0}^{m_n}\binom{m_n}{\ell}\mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] \le O(n^{-1}\log n)+\log\sup_{\beta\in[0,1]}\sup_{\gamma\in[0,1]}S_{m_n/n}(\beta,\gamma). \qquad (4.13)$$

For any $B\ge 0$, routine calculus (with a separate argument for $B=0$) shows that

$$\sup_{\beta\in[0,1]}\left(\frac{B^\beta}{\beta^\beta(1-\beta)^{1-\beta}}\right)=B+1,$$

with the supremum attained at $\beta=B/(1+B)$, so that from (2.6) and (4.13) we have

$$n^{-1}\log\mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] \le O(n^{-1}\log n)+\log\sup_{\gamma\in[0,1]}\left(\frac{(1+|\rho(1-2\gamma)|)^{m_n/n}}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right). \qquad (4.14)$$

Considering the transformation $\gamma\mapsto 1-\gamma$, we see that

$$\sup_{\gamma\in[1/2,1]}\left(\frac{(1+|\rho(1-2\gamma)|)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right)=\sup_{\gamma\in[0,1/2]}\left(\frac{(1+|\rho(2\gamma-1)|)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right)\le\sup_{\gamma\in[0,1/2]}\left(\frac{(1+\rho(1-2\gamma))^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right),$$

since, for $\gamma\in[0,1/2]$, $|\rho(2\gamma-1)|\le\rho(1-2\gamma)$, by Lemma B.1(iii). Hence from (4.14) we have, with $F_\rho(\alpha)$ as defined at (2.2), $n^{-1}\log\mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)]\le O(n^{-1}\log n)+F_\rho(m_n/n)$. Since $m_n/n\to\alpha$ and $\alpha\mapsto F_\rho(\alpha)$ is continuous (see Lemma 4.1),

$$\limsup_{n\to\infty} n^{-1}\log\mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)]\le F_\rho(\alpha).$$

For the lower bound, we use the fact that, by (2.6) and (3.1),

$$\mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)] \ge \sum_{\ell\in[m_n]\cap 2\mathbb{Z}}\binom{m_n}{\ell}\sum_{j=0}^{\lfloor n/2\rfloor}2^{-n}\binom{n}{j}(\rho(1-(2j/n)))^\ell$$
$$\ge 2^{-n}\sup_{\beta\in[0,1]:\beta m_n\in 2\mathbb{Z}}\sup_{\gamma\in[0,1/2]}\binom{m_n}{\beta m_n}\binom{n}{\gamma n}|\rho(1-2\gamma)|^{\beta m_n},$$

using the nonnegativity of the appropriate terms for both inequalities. Using the lower bound in (B.2), similarly to above, we obtain that $\liminf_{n\to\infty} n^{-1}\log\mathbb{E}^{\mathrm{bin}}_{\rho_n}[\mathcal{N}(n,m_n)]\ge F_\rho(\alpha)$. Hence combining the upper and lower bounds, we obtain (4.11). □

Now we can prove our main result on the expected number of null vectors.

Proof of Theorem 2.1. Lemma 4.5 shows that (2.3) holds for the case where $W_n=W_n^{\mathrm{bin}}$, and Lemma 4.2 shows that the result carries over to the general case. For the final

statement of the theorem, suppose that $\alpha < \alpha^*_\rho$ and that $\mathbb{P}[r_0 \le W_n \le r_1] = 1$ for some $r_0 \ge 3$ and $r_1 < \infty$. Lemma 4.3 shows that, for a suitable $\delta > 0$,

$$\sum_{\ell=1}^{\delta n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] = O(n^{2-r_0}). \tag{4.15}$$

For $\ell \ge \delta n$, we first restrict to the binomial model. Choose $\varepsilon > 0$ so that $(3\varepsilon)^\delta < 2^{-2\alpha}$. By a similar argument to (4.12), but splitting the supremum over j into two parts,

$$\begin{aligned}\sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] &\le (m_n+1)(n+1)2^{-n} \sup_{\delta n \le \ell \le m_n} \sup_{j:|j-(n/2)|\le \varepsilon n} \binom{m_n}{\ell}\binom{n}{j} |\rho(1-(2j/n))|^\ell \\ &\quad + (m_n+1)(n+1)2^{-n} \sup_{0 \le \ell \le m_n} \sup_{j:|j-(n/2)|> \varepsilon n} \binom{m_n}{\ell}\binom{n}{j} |\rho(1-(2j/n))|^\ell. \end{aligned} \tag{4.16}$$

Similarly to (4.14), the second term on the right-hand side of (4.16) is bounded above by

$$\exp\left\{ n \left(o(1) + \sup_{\gamma \in [0,(1/2)-\varepsilon]} F_{\rho, m_n/n}(\gamma) \right) \right\},$$

which decays to 0 exponentially fast, by Lemma 4.1(ii), since $m_n/n \to \alpha \in (0, \alpha^*_\rho)$. On the other hand, for $|j-(n/2)| \le \varepsilon n$, we have from Lemma B.1(i) that $|\rho(1-(2j/n))| \le 3\varepsilon$, for $\varepsilon > 0$ small enough, so that, since $\ell \ge \delta n$, $|\rho(1-(2j/n))|^\ell \le (3\varepsilon)^{\delta n}$, so the first term in the right hand side of (4.16) is bounded above by $(m_n+1)(n+1)2^{m_n}(3\varepsilon)^{\delta n}$. Hence

$$\limsup_{n\to\infty} n^{-1} \log \sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] < 0. \tag{4.17}$$

Next we use Lemma 3.5 to deduce a version of (4.17) with $\mathbb{P}_{\rho_n}$ in place of $\mathbb{P}^{\mathrm{bin}}_{\rho_n}$. Observe first that the $\mathbb{P}_{\rho_n}$-analogue of the sum in (4.17) consists of $O(n)$ nonnegative terms, so is bounded between the largest term and $O(n)$ times that same term, so that

$$n^{-1} \log \sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] = n^{-1} \log \sup_{\delta n \le \ell \le m_n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] + O(n^{-1}\log n).$$

By Lemma 3.5, for any $\varepsilon > 0$, there exist $\varepsilon_{n,\ell}$ with $|\varepsilon_{n,\ell}| \le \varepsilon$, uniformly for ℓ with $\delta n \le \ell \le m_n$ and n sufficiently large, such that

$$\binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] = \binom{m_n}{\ell} \mathbb{P}^{\mathrm{bin}}_{\rho_n}[A(n,\ell)] \exp\{\varepsilon_{n,\ell} n\}.$$

So now using (4.17) and the fact that $\varepsilon > 0$ was arbitary, we obtain

$$\limsup_{n\to\infty} n^{-1} \log \sum_{\ell=\delta n}^{m_n} \binom{m_n}{\ell} \mathbb{P}_{\rho_n}[A(n,\ell)] < 0,$$

which, combined with (4.15), yields (2.4). □

Remark 4.6. *Write $\beta_0 = \frac{\rho(1-2\gamma_0)}{1+\rho(1-2\gamma_0)} \in [0,1/2]$, where $\gamma_0 = \gamma_0(\alpha) \in [0,1/2]$ is the value of γ for which the supremum in* (2.2) *is attained. The proofs above show that for $\alpha > \alpha^*_\rho$ the exponential rate in* (2.3) *is dominated by null vectors using proportion $\beta_0 = \beta_0(\alpha)$ of the (roughly αn) available rows (and also possibly those using proportion $1-\beta_0$ of the rows, due to parity effects).*

5 Cores of sparse random hypergraphs

5.1 Hypercycles and 2-cores

Recall the definitions from Section 2.3. The connection between the 2-core and hypercycles was exploited by Cooper [5, p. 371], following an idea that he attributes to Molloy (see [4, p. 268]). The connection is demonstrated by the following result.

Lemma 5.1. *Suppose that the 2-core $\mathcal{C} := Core(\mathcal{E})$ of a hypergraph $(\mathcal{V}, \mathcal{E})$ has vertex span $V(\mathcal{C}) \subseteq \mathcal{V}$ and size (number of hyperedges) $|\mathcal{C}|$.*

(i) Any hyperedge $E \notin \mathcal{C}$ cannot belong to a hypercycle of $(\mathcal{V}, \mathcal{E})$.

(ii) If $\mathcal{C} = \emptyset$, then $(\mathcal{V}, \mathcal{E})$ contains no hypercycle.

(iii) If $|V(\mathcal{C})| < |\mathcal{C}|$, then $(\mathcal{V}, \mathcal{E})$ contains a hypercycle.

Proof. If there are s hyperedges not in the 2-core $\mathcal{C}$, there exists a labelling of them as $E_1, E_2, \ldots, E_s$ with the property that, for every j, E_j has some vertex with degree one after hyperedges $E_1, E_2, \ldots, E_{j-1}$ are removed. Suppose $(\mathcal{V}, \mathcal{E})$ has some hypercycle $\mathcal{F} \neq \emptyset$. None of $E_1, E_2, \ldots, E_s$ can belong to $\mathcal{F}$: otherwise, there would be some minimum j for which $E_j \in \mathcal{F}$, and this E_j has some vertex v of degree one in the partial hypergraph from which $E_1, E_2, \ldots, E_{j-1}$ have been removed, which contains $\mathcal{F}$; so v cannot have even degree in $\mathcal{F}$, which is a contradiction. This proves (i), and (ii) follows. For (iii), say $c := |V(\mathcal{C})| < |\mathcal{C}| =: r$. Then there are $2^r - 1$ non-empty partial hypergraphs, but only $2^c < 2^r - 1$ possible indicator vectors for a set of vertices of odd degree. By the pigeonhole principle, there must be two distinct partial hypergraphs $\mathcal{F}, \mathcal{F}' \subseteq \mathcal{E}$ for which the sets of vertices of odd degree are the same. Then $\mathcal{F} \triangle \mathcal{F}'$ is a hypercycle. □

5.2 The 2-core in uniform random hypergraphs

In this section we consider a certain *uniform random* hypergraph model, which is different from (but related to) the hypergraph model induced by our random matrix $M(n, m_n)$; in Section 5.3 we will connect the two models.

Recall from Section 2.3 that we may represent a hypergraph by an incidence matrix, and the degree of a vertex is the number of incidences (non-zero entries) in the corresponding column of the incidence matrix. The *weight* of a hyperedge is its number of incident vertices, i.e., the number of incidences in the corresponding row of the matrix.

A natural probability model for a random hypergraph is to fix the multiset of vertex degrees and the multiset of hyperedge weights in advance (subject to a consistency condition), and sample uniformly from the hypergraphs with these collections of vertex degrees and hyperedge weights. This gives a *uniform random hypergraph*.

Darling and Norris [9] analyse the statistical properties of the 2-cores for sequences of uniform random hypergraphs, assuming a uniform bound on the hyperedge weights and vertex degrees. In unpublished work of the same authors, the uniform bounds are replaced by third moments assumptions. For the present paper, we require a more modest relaxation of the conditions of [9], to cover the case where the row weights remain uniformly bounded but the vertex degrees are approximately Poisson distributed.

For each n, define vectors of nonnegative integers $\mathbf{d}_n := (d_n(k) : k \in \mathbb{Z}_+)$ and $\mathbf{w}_n := (w_n(k) : k \in \mathbb{N})$ with $\sum_{k\geq 0} d_n(k) = n$ and $m_n := \sum_{k\geq 1} w_n(k)$; we assume that $\mathbf{d}_n$ and $\mathbf{w}_n$ are compatible in the sense that $\sum_{k\geq 1} k w_n(k) = \sum_{k\geq 0} k d_n(k) < \infty$. We also assume that $m_n \to \infty$. Suppose that for each $i \in \mathbb{N}$ and $j \in \mathbb{Z}_+$,

$$\lim_{n\to\infty} \frac{w_n(i)}{\sum_{k\geq 1} w_n(k)} = \rho_i; \quad \lim_{n\to\infty} \frac{d_n(j)}{n} = \nu_j. \tag{5.1}$$

Define generating functions $\rho(s) := \sum_{k\geq 1} \rho_k s^k$ and $\nu(s) := \sum_{k\geq 0} \nu_k s^k$. We assume that the weights are uniformly bounded and the degree distribution has all moments, so that the means $\rho'(1)$ and $\nu'(1)$ corresponding to these distributions are finite.

Consider a sequence of random hypergraphs with n vertices and m_n hyperedges, selected uniformly from those hypergraphs with edge weight multiplicities $\mathbf{w}_n$ and vertex degree multiplicities $\mathbf{d}_n$. Let (E_n, v_n) be a *random incidence*, sampled uniformly at random from all incidences in the nth hypergraph. Denote the weight of E_n by $1 + S_n$, and the degree of v_n by $1 + L_n$; thus S_n counts the other vertices in this hyperedge, and L_n counts the other hyperedges incident to this vertex. Size bias occurs here: the probability that E_n has weight k is proportional to k times the number of rows of weight k, and similarly the probability that the degree of v_n is d is proportional to d times the number of degree d vertices. Given ρ_w and ν_d describing via (5.1) the limiting weight and degree distributions, we may thus compute a pair of limiting probability generating functions:

$$\lambda(s) := \lim_{n\to\infty} \mathbb{E}[s^{L_n}] =: \sum_{d=0}^{\infty} \lambda_d s^d; \quad \sigma(s) := \lim_{n\to\infty} \mathbb{E}[s^{S_n}] =: \sum_{w=0}^{\infty} \sigma_w s^w, \tag{5.2}$$

where, due to the size biasing, the coefficients in (5.2) are given by

$$\lambda_d = \frac{(d+1)\nu_{d+1}}{\nu'(1)}; \quad \sigma_w = \frac{(w+1)\rho_{w+1}}{\rho'(1)}.$$

Hence the generating functions themselves become:

$$\lambda(s) = \frac{\nu'(s)}{\nu'(1)}; \quad \sigma(s) = \frac{\rho'(s)}{\rho'(1)}. \tag{5.3}$$

To avoid triviality, we assume that $\sigma_0 = 0$ (equivalently, $\rho_1 = 0$), i.e., there are no 1-edges, and $\lambda_0 \notin \{0, 1\}$ (otherwise the 2-core is of no interest). Define

$$\varphi(s) := 1 - \lambda(1 - \sigma(s)). \tag{5.4}$$

Since $\sigma_0 \neq 1$ and $\lambda_0 \neq 1$, we deduce that $\varphi : [0, 1] \to \mathbb{R}$ is strictly increasing; moreover $\varphi(0) = 1 - \lambda(1 - \sigma(0)) = 0$ (since $\sigma_0 = 0$) and $\varphi(1) = 1 - \lambda_0 \in (0, 1)$, so φ takes values in $[0, 1)$, and there exists a largest solution g^* in $[0, 1)$ of the equation $\varphi(s) = s$. That is,

$$g^* := \sup\{s \in [0, 1) : \varphi(s) = s\}. \tag{5.5}$$

In the case where $g^* > 0$ and the curve $y = \varphi(s)$ crosses the curve $y = s$ (rather than just touching it) at $s = g^*$, we also have

$$g^* = \sup\{s \in (0, 1) : \varphi(s) > s\}. \tag{5.6}$$

Now we can state the result on the 2-core that we shall use, which amounts to a variant of Theorem 7.1 of [9].

Theorem 5.2. *Consider a sequence of uniform random hypergraphs associated with sequences $\mathbf{w}_n$ and $\mathbf{d}_n$ satisfying* (5.1) *with $\rho_w = 0$ for all w large enough and $\sum_{d\geq 1} \nu_d d^\beta < \infty$ for all $\beta > 0$. Suppose that the corresponding pair* (5.2) *of random-incidence generating functions has $\sigma_0 = 0$, $\lambda_0 \notin \{0, 1\}$, and is such that g^*, given by* (5.5)*, has either $g^* = 0$ or g^* satisfying* (5.6)*. Then the following hold a.s. in the limit as $n \to \infty$.*

(i) If $g^* = 0$, the proportion of hyperedges which survive in the 2-core converges to zero.

(ii) If $g^* > 0$, then for any $k \in \mathbb{Z}_+$ with $\rho_k > 0$, the proportion of weight-k hyperedges which survive in the 2-core is asymptotically $(g^*)^k$; overall, a proportion $\rho(g^*)$ of hyperedges survive, and a proportion $g^*\sigma(g^*)$ of incidences.

(iii) If $g^* > 0$, then for any $d, k \in \mathbb{N}$ with $2 \le d \le k$ and $\nu_k > 0$, the proportion of vertices of degree k whose degree in the 2-core is d converges to

$$\binom{k}{d}\sigma(g^*)^d(1-\sigma(g^*))^{k-d}.$$

(iv) If $g^* > 0$, the 2-core is again a uniform random hypergraph, given its hyperedge weights and vertex degrees, whose distributions are determined by the previous assertions.

As mentioned above, in [9] all but finitely many coefficients of the generating functions (5.2) were taken to be zero, but the methods admit the modest extension of this section, and indeed can be extended to the case where $\lambda''(1)$ and $\sigma''(1)$ are finite, corresponding to finite third moments for hyperedge weight and vertex degree distributions. Because of its proximity to the result in [9], we do not prove Theorem 5.2 here.

5.3 Application to the random matrix model

We return to the random matrix model used in the rest of the paper, so our random incidence matrix will be $M(n, m_n)$ described in Section 2.1, i.e., with i.i.d. rows with W_n-distributed weights, and corresponding generating function $\rho_n(s)$ having limit $\rho(s)$. To justify application of Theorem 5.2 in this setting, we give the following strong laws of large numbers for the empirical distributions of the row and column weights of M.

Lemma 5.3. *Suppose $m_n \in \mathbb{N}$ with $m_n/n \to \alpha$ as $n \to \infty$, with $\alpha > 0$, and the W_n are uniformly bounded. For $k \in \mathbb{Z}_+$, let $N_k(n)$ be the number of rows of $M(n, m_n)$ of weight k, and let $\tilde{N}_k(n)$ be the number of columns of $M(n, m_n)$ of degree k. Then a.s.,*

$$\lim_{n\to\infty} m_n^{-1} N_k(n) = \mathbb{P}[W = k], \quad \textit{and} \tag{5.7}$$

$$\lim_{n\to\infty} n^{-1}\tilde{N}_k(n) = \frac{\mathrm{e}^{-\mu}\mu^k}{k!}, \tag{5.8}$$

where we set $\mu := \alpha\mathbb{E}[W] = \alpha\rho'(1)$. Moreover, the total number of incidences satisfies

$$\lim_{n\to\infty} n^{-1}\sum_{k\ge1} kN_k(n) = \lim_{n\to\infty} n^{-1}\sum_{k\ge1} k\tilde{N}_k(n) = \mu, \text{ a.s.} \tag{5.9}$$

Proof. First note that ${m_n}^{-1}\mathbb{E}[N_k(n)] = \mathbb{P}[W_n = k]$, which converges to $\mathbb{P}[W = k]$ by assumption. To deduce almost sure convergence from this convergence in means, we use the Azuma–Hoeffding inequality in a standard way, as follows. Fix n and for $1 \le i \le m_n$ let $\mathcal{F}_i$ be the σ-algebra generated by the rows $X_1, \dots, X_i$ of $M(n, m_n)$. Define $\xi_i = \mathbb{E}[N_k(n) \mid \mathcal{F}_i]$, with $\xi_0 = \mathbb{E}[N_k(n)]$. Since resampling a single row changes the number of rows of weight k by at most 1, we have for $1 \le i \le m_n$ that

$$|\xi_i - \xi_{i-1}| = |\mathbb{E}[N_k(n) - N_k(n,i) \mid \mathcal{F}_i]| \le 1,$$

where $N_k(n,i)$ is defined like $N_k(n)$ but based on a matrix with row i resampled. Let $\varepsilon > 0$. The Azuma–Hoeffding inequality applied to the martingale $(\xi_0, \dots, \xi_{m_n})$ yields

$$\mathbb{P}[|N_k(n) - \mathbb{E}[N_k(n)]| > \varepsilon n] \le 2\exp(-\varepsilon^2 n/(4\alpha)), \text{for all } n \text{ large enough},$$

so by the first Borel–Cantelli lemma, we have $|N_k(n) - \mathbb{E}[N_k(n)]| \le \varepsilon n$ for all but finitely many n almost surely. Combined with the convergence of the mean, this gives us (5.7).

The remaining two parts of the lemma use the assumption $\mathbb{P}[W \le r_1] = 1$ for $r_1 < \infty$. To prove (5.8) note that the degree of the first column of $M(n, m_n)$ is binomially distributed with parameters m_n and $\mathbb{E}[W_n]/n$. Hence $\mathbb{E}[\tilde{N}_k(n)/n] = \mathbb{P}[\mathrm{Bin}(m_n, \mathbb{E}[W_n]/n) = k]$, which tends to $\mathrm{e}^{-\mu}\mu^k/k!$ as $n \to \infty$ by binomial-Poisson convergence. Given convergence of means, we may prove (5.8) by a similar Azuma–Hoeffding argument to that for (5.7), since resampling a single row changes the number of columns of degree k by at most r_1.

For the final statement in the lemma, we have that

$$n^{-1}\sum_{k\ge 1} kN_k(n) = (m_n/n)\sum_{k=1}^{r_1} km_n^{-1}N_k(n) \to \alpha\sum_{k=1}^{r_1} k\mathbb{P}[W=k], \text{ a.s.},$$

by (5.7), and then (5.9) follows. □

Corollary 5.4. *Consider the random matrix model $M(n, m_n)$ with $W_n \xrightarrow{d} W$, where the W_n are uniformly bounded. Suppose that $m_n/n \to \alpha > 0$. Then, a.s., taken as hypergraph incidence matrices the sequence $M(n, m_n)$ defines a sequence of uniform random hypergraphs whose row weight and vertex degree distributions satisfy* (5.1) *with ρ_w and ν_d given by $\rho_w = \mathbb{P}[W = k]$ and $\nu_d = \mathrm{e}^{-\mu}\mu^d/d!$ respectively, where $\mu := \alpha\mathbb{E}[W]$.*

Proof. Since the distribution of $M(n, m)$ is invariant under permutations of the rows or columns, conditional on the empirical distribution of row and column weights, all possible outcomes with those row and column weight distributions are equally likely, so this conditional distribution is indeed uniform. Moreover by Lemma 5.3 the limiting proportion of rows of weight k is given by $\mathbb{P}[W = k]$ and the limiting proportion of columns of degree k is $\mathbb{P}[D = k]$ where $D \sim \mathrm{Po}(\mu)$ is Poisson with mean $\mu := \alpha\rho'(1)$. Hence conditionally on this sequence of empirical distributions, almost surely we have a sequence of random matrices satisfying the hypotheses of Section 5.2. □

In the notation of Section 5.2, in this case $\nu(s) = \sum_{d=0}^{\infty} \mathrm{e}^{-\mu}\frac{(s\mu)^d}{d!} = \mathrm{e}^{\mu(s-1)}$ is the generating function of a $\mathrm{Po}(\mu)$ random variable, so, by (5.3), the pair (5.2) becomes

$$\lambda(s) = \mathrm{e}^{\mu(s-1)}; \quad \sigma(s) = \frac{\rho'(s)}{\rho'(1)}.$$

In this case, we have from (5.4) that

$$\varphi(s) := 1 - \lambda(1-\sigma(s)) = 1 - \mathrm{e}^{-\mu\sigma(s)} = 1 - \mathrm{e}^{-\alpha\rho'(s)},$$

and to emphasize the dependence on α we will use the notation φ_α for φ from now on. Recall from (5.5) that g^* was defined as the largest $s \in [0, 1)$ for which $\varphi_\alpha(s) = s$. In order for the model of this section to fit into the setting discussed in Section 5.2, we need to assume that $\sigma_0 = 0$ and $\lambda_0 \notin \{0, 1\}$. Here $\lambda_0 = \mathrm{e}^{-\mu} = \mathrm{e}^{-\alpha\mathbb{E}[W]}$ and $\sigma_0 = \frac{\mathbb{P}[W=1]}{\mathbb{E}[W]}$. So it suffices to assume that $\alpha > 0$, $\mathbb{P}[W \ge 2] = 1$, and $\mathbb{E}[W] < \infty$; in this case the argument in Section 5.2 shows that g^* is well defined.

Note that g^* depends both on ρ and on α; in this section we write $g^* = g^*(\alpha)$ to emphasize the dependence on α; we will show (see Lemma 5.5) that the present definition is equivalent to that at (2.11) given in Section 2.2. For any solution $s \in [0, 1)$ to $\varphi_\alpha(s) = s$, so in particular for $s = g^*(\alpha)$, provided $\rho'(s) \ne 0$, we have $\alpha = h(s)$ as given by (2.9).

We note some facts about $g^*(\alpha)$; recall the definition of $\alpha^\sharp_\rho$ from (2.10).

Lemma 5.5. *Suppose that* $\mathbb{P}[W \geq 2] = 1$ *and* $\mathbb{E}[W] < \infty$. *With the convention* $\sup \emptyset = 0$, *the definition* (2.11) *is equivalent to the definition* (5.5) *of* $g^*(\alpha)$ *as the largest solution of* $\varphi_\alpha(s) = s$. *Also,* $\alpha_\rho^\sharp \in [0,1]$, *and* $g^*(\alpha) = 0$ *for all* $\alpha \in [0, \alpha_\rho^\sharp)$, *and for* $\alpha > \alpha_\rho^\sharp$, *the function* $g^*(\alpha)$ *is positive and strictly increasing, with* $g^*(\alpha) \uparrow 1$ *as* $\alpha \to \infty$.

Now assume also that $\mathbb{P}[W \geq 3] = 1$ *and* $\mathbb{E}[W^2] < \infty$. *Then the following hold.*

(i) We have $g^*(\alpha_\rho^\sharp) \in (0,1)$ *and* $\alpha_\rho^\sharp = h(g^*(\alpha_\rho^\sharp)) \in (0, \infty)$.

(ii) The function g^* *is right continuous, and there is a finite set* $\mathcal{D}_\rho \subset (0,\infty)$, *with* $\alpha_\rho^\sharp = \inf \mathcal{D}_\rho$, *such that* g^* *is continuous apart from jumps at points of* $\mathcal{D}_\rho$. *For each* $\alpha \in \mathcal{D}_\rho$, $h(g^*(\alpha)) = \alpha$ *is a local minimum for* h.

(iii) If $\alpha \notin \mathcal{D}_\rho$, *then* $g^*(\alpha)$ *satisfies the crossing condition* (5.6).

Proof. Since $\mathbb{P}[W \geq 2] = 1$ and $\mathbb{E}[W] < \infty$, we have $\varphi_\alpha(1) < 1$ and $\varphi_\alpha(0) = 0$. Therefore by continuity we may rewrite (5.5) using the convention $\sup \emptyset = 0$ as

$$g^*(\alpha) = \sup\{s \in (0,1) : \varphi_\alpha(s) \geq s\}.$$

By the definition (2.9) of h, for $s \in (0,1)$ it is easy to check that $\varphi_\alpha(s) \geq s$ if and only if $h(s) \leq \alpha$, and this shows that (2.11) and (5.5) give equivalent definitions of $g^*(\alpha)$.

By (2.9) and subsequent remarks, $h(x)$ is positive, continuous in x, and tends to infinity as $x \uparrow 1$. By the definition (2.10), and the subsequent remark, $\alpha_\rho^\sharp \in [0,1]$. By the definition (2.11), it is clear that $g^*(\alpha) = 0$ for $\alpha \in [0, \alpha_\rho^\sharp)$, and the fact that $g^*(\alpha)$ is positive and strictly increasing for $\alpha \in (\alpha_\rho^\sharp, \infty)$ is easily deduced from the continuity of h. Also, given $\varepsilon \in (0,1)$ we can choose α with $h(1-\varepsilon) \leq \alpha$ so that $g^*(\alpha) \geq 1 - \varepsilon$, and together with the monotonicity of g^* this shows $g^*(\alpha) \to 1$ as $\alpha \to \infty$.

For part (i), under the extra assumption $\mathbb{P}[W \geq 3] = 1$ we have h going to infinity at 0 and at 1, and by continuity h attains its infimum on $(0,1)$, so using (2.10) and (2.11) we have that $g^*(\alpha_\rho^\sharp)$ is the supremum of a non-empty compact set contained in $(0,1)$, and so lies in (0,1). The last part of (i) also follows from the continuity of h.

For part (ii), under the extra assumption $\mathbb{E}[W^2] < \infty$, note first that if $0 \leq y < \alpha_\rho^\sharp$ then $g^*(y) = 0$. Hence g^* is continuous at y for all $y < \alpha_\rho^\sharp$.

Now let $y \geq \alpha_\rho^\sharp$; note that by (2.11) and continuity of h, we have $h(g^*(y)) = y$. Take a monotonic sequence y_n tending to y; set $x_n = g^*(y_n)$.

Suppose first that $y_n \downarrow y$. Then x_n is nonincreasing; denoting the limit by x_∞ we have $h(x_n) = y_n$ so $h(x_\infty) = y$ by continuity, and therefore $x_\infty \leq g^*(y)$ by (2.9). Since also $x_n \geq g^*(y)$ by monotonicity we have $x_\infty = g^*(y)$; hence g^* is right-continuous at y.

Now suppose instead that $y_n \uparrow y$. Set $x = g^*(y)$. If h does not have a local minimum at x then $\liminf g^*(y_n) \geq x$, so that $x_n \to x$, and hence g^* is left-continuous at y. Hence, if g^* is discontinuous at y then h has a local minimum at $g^*(y)$.

The function h' is analytic and non-constant on $(0,1)$ so its zeros do not accumulate except possibly at 0 or 1. But $h'(x) = 0$ if and only if $\rho'(x)/\rho''(x) = -(1-x)\log(1-x)$. This equality yields a contradiction as $x \uparrow 1$ (since $\mathbb{E}[W^2] < \infty$ and $\mathbb{E}[W] > 0$) and as $x \downarrow 0$ since, if $r_0 \geq 3$ is the smallest possible value of W,

$$\lim_{x \to 0} \left(-\frac{\rho''(x)}{\rho'(x)} (1-x)\log(1-x) \right) = r_0 - 1 > 1.$$

Hence there exists $\varepsilon > 0$ such that $h'(x) \neq 0$ for $1 - \varepsilon < x < 1$ and for $0 < x < \varepsilon$, so h has only finitely many local minima in $(0,1)$. Thus h has a local minimum at $g^*(y)$ for at most finitely many y. This completes the proof of (ii).

For part (iii) note that, for $s \in (0,1)$, $\varphi_\alpha(s) > s$ if and only if $h(s) < \alpha$, so (5.6) reads $g^*(\alpha) = \sup\{s \in (0,1) : h(s) < \alpha\}$, which for $\alpha \notin \mathcal{D}_\rho$ agrees with (2.11). □

To apply the results in Section 5.2 to $M(n, m_n)$, we need to assume that $g^*(\alpha)$ either is zero or satisfies (5.6). Lemma 5.5 shows it suffices to take $\alpha \notin \mathcal{D}_\rho$.

By Theorem 5.2(ii) and (5.9), n^{-1} times the number of incidences which survive in the 2-core converges a.s. to

$$\mu g^* \sigma(g^*) = (\alpha\rho'(1))g^* \frac{\rho'(g^*)}{\rho'(1)} = \alpha g^* \rho'(g^*) = -g^* \log(1 - g^*). \tag{5.10}$$

By Theorem 5.2(iii) and (5.8), for $d \geq 2$ the proportion of original vertices whose degree in the 2-core is d is asymptotically

$$\begin{aligned}\sum_{k \geq d} \mathrm{e}^{-\mu} \frac{\mu^k}{k!} \binom{k}{d} \sigma(g^*)^d (1 - \sigma(g^*))^{k-d} &= \mathrm{e}^{-\mu} \frac{(\mu\sigma(g^*))^d}{d!} \sum_{j \geq 0} \frac{\mu^j (1 - \sigma(g^*))^j}{j!} \\ &= \mathrm{e}^{-\mu\sigma(g^*)} \frac{(\mu\sigma(g^*))^d}{d!},\end{aligned} \tag{5.11}$$

the remainder having degree 0 in the 2-core (the algorithm of Section 5.1 never deletes any columns). In other words, the 2-core vertex degrees have the distribution of a random variable $D\mathbf{1}\{D \neq 1\}$, where $D \sim \mathrm{Po}(\mu\sigma(g^*))$; by (5.10), $\mu\sigma(g^*) = \alpha\rho'(g^*)$. As a check on the previous calculation of the number of surviving incidences, n^{-1} times the total number of incidences in the 2-core should converge to the mean of the vertex-degree distribution, which is $\alpha\rho'(g^*)(1 - \mathrm{e}^{-\alpha\rho'(g^*)}) = \alpha g^* \rho'(g^*)$, as in (5.10).

5.4 Proofs of Theorems 2.2 and 2.4

We are now in a position to present the proof of Theorem 2.4.

Proof of Theorem 2.4. By Corollary 5.4, a.s. our sequence of random matrices satisfies the hypotheses of Theorem 5.2. If $\alpha < \alpha_\rho^\sharp$, then $g^*(\alpha) = 0$, and Theorem 5.2 implies the 2-core has $o(n)$ rows. From now on suppose $\alpha > \alpha_\rho^\sharp$, so $g^* = g^*(\alpha) > 0$ by Lemma 5.5.

For the statement (i), note that out of $m_n \sim \alpha n$ rows, a proportion $\rho(g^*)$ survives, by Theorem 5.2(ii). For (ii), the discussion around (5.11) implies that the proportion of the n original vertices whose degree in the 2-core is non-zero is obtained by subtracting from 1 the mass that a $\mathrm{Po}(\nu)$ random variable places on $\{0, 1\}$.

For (iii), we compare the limits in (i) and (ii). Suppose that these limits satisfy

$$\alpha\rho(g^*) > 1 - \mathrm{e}^{-\alpha\rho'(g^*)}(1 + \alpha\rho'(g^*)). \tag{5.12}$$

By our assumptions on α and W, we have $\rho(0) = \rho'(0) = 0$ and $g^* > 0$, which implies that $\rho(g^*)$ and $\rho'(g^*)$ are both positive. Then we may rewrite (5.12) as

$$\begin{aligned}\alpha\rho(g^*) &> 1 - (1 - g^*)(1 + \alpha\rho'(g^*)) \\ &= g^* + (1 - g^*)\log(1 - g^*),\end{aligned}$$

using the definition of g^*. Now substituting in $\alpha = h(g^*(\alpha))$ for α on the left-hand side of the last display (given $\rho'(g^*) > 0$) we may rewrite the last inequality as $\psi(g^*) < 0$, where ψ is defined by (2.8). Similarly, $\psi(g^*) > 0$ is equivalent to

$$\alpha\rho(g^*) < 1 - \mathrm{e}^{-\alpha\rho'(g^*)}(1 + \alpha\rho'(g^*)). \tag{5.13}$$

If $\psi(g^*) < 0$, then (5.12) holds and the limit in (i) in strictly greater than the limit in (ii), which shows that the 2-core eventually has more rows than occupied columns, and vice versa if $\psi(g^*) > 0$ (so that (5.13) holds).

Recall the definition of $\underline{\alpha}_\rho$ from (2.12). By Lemma 5.5(iii), g^* has finitely many discontinuities. Either $\underline{\alpha}_\rho$ is a continuity point for $g^*(\,\cdot\,)$ (and hence for $\psi(g^*(\,\cdot\,))$), or else $\underline{\alpha}_\rho \in \mathcal{D}_\rho$ with $\psi(g^*(\underline{\alpha}_\rho)) < 0$ and no other point of $\mathcal{D}_\rho$ is in a neighbourhood of $\underline{\alpha}_\rho$. In either case, $\psi(g^*(\,\cdot\,)) < 0$ over an interval of the form $(\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$ with $\delta > 0$. □

We next state a result giving an upper bound for $\underline{\alpha}_\rho$.

Proposition 5.6. *Suppose that* $\mathbb{P}[W \geq 3] = 1$ *and* $\mathbb{E}[W^2] < \infty$. *Then* $\underline{\alpha}_\rho \leq 1$.

Proof. We know from Lemma 5.5 that $\alpha_\rho^\sharp \leq 1$, so if $\underline{\alpha}_\rho \leq \alpha_\rho^\sharp$ there is nothing to prove. Hence we assume $\underline{\alpha}_\rho > \alpha_\rho^\sharp$ from now on. First we show that

$$\text{for any } \varepsilon > 0 \text{ there exists } \alpha \in (\underline{\alpha}_\rho - \varepsilon, \underline{\alpha}_\rho), \text{ such that } \psi(g^*(\alpha)) > 0. \tag{5.14}$$

By the definition (2.12) of $\underline{\alpha}_\rho$, and the assumption $\underline{\alpha}_\rho > \alpha_\rho^\sharp$, if (5.14) fails then there exists $\delta > 0$ such that $\psi \circ g^*$ is identically zero on the interval $I := (\underline{\alpha}_\rho - \delta, \underline{\alpha}_\rho)$, and by taking δ small enough we may assume the interval I contains no discontinuities of g^*. But then the image $J := g^*(I)$ is also an open interval because g^* is continuous and strictly increasing on I. So we would then have ψ identically zero on J, which would contradict the fact that ψ is analytic and non-constant on $(0, 1)$. Thus (5.14) must hold as asserted.

Observe next that every time the 2-core algorithm deletes a row, it has to create at least one column of degree zero, and possibly more. So the aspect ratio (i.e., number of rows divided by number of occupied columns) is nondecreasing at each step of the algorithm, provided the initial aspect ratio is at least 1. Hence the aspect ratio of a non-empty 2-core is at least as large as the aspect ratio of the original incidence matrix to which the algorithm is applied, provided the latter is at least 1.

So if $m_n/n \to \alpha > 1$, the aspect ratio of the original matrix exceeds 1 for all n large enough, and hence so does the aspect ratio of any non-empty 2-core. Suppose that $\underline{\alpha}_\rho > 1$. By (5.14) and the finiteness of $\mathcal{D}_\rho$, there exists $\alpha' \in (1, \underline{\alpha}_\rho) \setminus \mathcal{D}_\rho$ such that $\psi(g^*(\alpha')) > 0$. Then, by Theorem 2.4(iii), with $m_n/n \to \alpha = \alpha'$, the 2-core has aspect ratio less than 1 for all n large enough, which contradicts the previous conclusion that $\alpha > 1$ implied the 2-core having limiting aspect ratio greater than 1. Hence $\underline{\alpha}_\rho \leq 1$. □

The situation in Theorem 2.4(iii) is clarified by the following facts on h and ψ.

Proposition 5.7. *Suppose that* $\mathbb{P}[W \geq 3] = 1$ *and* $\mathbb{E}[W^2] < \infty$. *Then* $0 < \alpha_\rho^\sharp \leq \underline{\alpha}_\rho \leq 1$. *The function* ψ *has at least one zero in* $(0, 1)$, *and* h *has at least one local minimum in* $(0, 1)$. *Suppose that the following condition holds:*

(a) h *has a single local minimum* $x_\rho^\sharp$ *in* $(0, 1)$, *with* $h(x_\rho^\sharp) = \inf_{x \in (0,1)} h(x)$.

Then $x_\rho^\sharp$ *is the location of the unique local maximum of* ψ *in* $(0, 1)$, $\psi(x_\rho^\sharp) > 0$, *and the interval* $(0, 1)$ *contains exactly one zero of* ψ, *denoted* x_ρ^*, *which satisfies* $x_\rho^\sharp < x_\rho^*$. *Moreover,* $\underline{\alpha}_\rho = h(x_\rho^*) > \alpha_\rho^\sharp$, *and*

$$\psi(g^*(\alpha)) \begin{cases} > 0 & \text{for all } \alpha \in (\alpha_\rho^\sharp, \underline{\alpha}_\rho) \\ < 0 & \text{for all } \alpha > \underline{\alpha}_\rho. \end{cases} \tag{5.15}$$

Finally, in the fixed row-weight case where where $W = r \geq 3$ *a.s., condition (a) holds, and the unique positive zero of* ψ *is* $x_r^* \in (\frac{r-2}{r-1}, 1)$.

An important observation that helps to explain the close connection between the functions h and ψ (apparent in Figure 1, for example) and will also form an ingredient in the proof of Proposition 5.7 is the following result.

Lemma 5.8. *For all* $x \in (0, 1)$, $\psi'(x)$ *has the same sign as* $-h'(x)$, *so, in particular, the locations of the local minima of* h *correspond exactly to the locations of the local maxima of* ψ *in* $(0, 1)$. *Moreover, if* $\mathbb{E}[W^2] < \infty$, *then as* $x \downarrow 0$ *we have*

$$\begin{aligned} \psi(1-x) &= 1 - h(1-x) - x - \frac{\mathbb{E}[W(W-1)]}{2\mathbb{E}[W]}(1 + o(1))x^2 \log x \\ &= 1 - h(1-x) - x + o(x). \end{aligned} \tag{5.16}$$

Proof. Differentiating (2.8), we obtain

$$\psi'(x) = -\frac{\rho(x)}{\rho'(x)}\left(\frac{1}{1-x} + \frac{\rho''(x)}{\rho'(x)}\log(1-x)\right). \tag{5.17}$$

On the other hand, from (2.9), we have that, for $x \in (0,1)$,

$$h'(x) = \frac{1}{\rho'(x)}\left(\frac{1}{1-x} + \frac{\rho''(x)}{\rho'(x)}\log(1-x)\right) = -\frac{1}{\rho(x)}\psi'(x),$$

by comparison with (5.17). Finally, (5.16) follows from a routine calculation. □

Before completing the proof of Proposition 5.7, we collect some further remarks and examples. The main complication in the interpretation of Theorem 2.4(iii) is due to discontinuity of g^*, so $\{\psi(g^*(\alpha)) : \alpha \geq 0\}$ is only a subset of $\{\psi(x) : x \in [0,1)\}$. Let

$$\mathcal{G}_\rho := \{g^*(\alpha) : \alpha \geq \alpha_\rho^\sharp\}.$$

By Lemma 5.5(ii), $\mathcal{G}_\rho$ is a union of finitely many intervals $\mathcal{G}_\rho = [g_1^-, g_1^+) \cup \cdots \cup [g_\ell^-, g_\ell^+)$ where $g_1^- < g_1^+ < g_2^- < \cdots < g_\ell^+$, and, for each k, $g_k^- = g^*(\alpha)$ for $\alpha \in \mathcal{D}_\rho$, and $h(g_k^-)$ is a local minimum. Recall that $\alpha = h(g^*(\alpha))$ and $\alpha \mapsto g^*(\alpha)$ is increasing for $\alpha > \alpha_\rho^\sharp$ (see Lemma 5.5), so $x \mapsto h(x)$ must be increasing on $\mathcal{G}_\rho$. So in fact $\alpha_\rho^\sharp = h(g_1^-) < \cdots < h(g_\ell^-)$. The 'curve' $\psi(g^*(\alpha))$, $\alpha \geq \alpha_\rho^\sharp$ is then a (discontinuous) trace of $\psi(x)$, where x runs over $\mathcal{G}_\rho$, piecewise continuously on intervals starting at g_k^- which, by Lemma 5.8, correspond to local *maxima* of ψ. Figures 1 and 2 give some illustrations of possible behaviour. Observe that $\psi(x)$ is not necessarily decreasing for all $x \in \mathcal{G}_\rho$.

Note that condition (a) in Proposition 5.7 is not necessary for the sharp transition property (5.15) to hold. Two other relevant conditions are:

(b) ψ has a single zero in $(0,1)$;
(c) the global minimum of h on $(0,1)$ is the rightmost local minimum.

If $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W] < \infty$, then $h(x) \to \infty$ as $x \to 0$ and as $x \to 1$, so (a) $\Rightarrow$ (c), while in the course of the proof of Proposition 5.7 below, we show that (a) $\Rightarrow$ (b) as well. We mention 3 illustrative examples.

- An example for which conditions (a) and (b) do not hold but (c) does is provided by $\rho(s) = 0.9s^3 + 0.1s^{38}$, for which ψ has 3 positive zeros (see Figure 3).
- An example for which condition (b) holds but conditions (a) and (c) do not is $\rho(s) = 0.9s^3 + 0.1s^{24}$, for which g^* has two discontinuities (see Figure 1).
- An example in which none of (a), (b) or (c) holds and where (5.15) fails is provided by $\rho(s) = 0.9183s^3 + 0.04s^{19} + 0.0417s^{41}$ (see Figure 2).

Proof of Proposition 5.7. First we show that if $\mathbb{P}[W \geq 3] = 1$ and $\mathbb{E}[W] < \infty$, then ψ has at least one zero in $(0,1)$. So suppose that there exists an integer $r \geq 3$ for which $\mathbb{P}[W \geq r] = 1$ and $\mathbb{P}[W = r] = p > 0$. Then $\rho(s) \sim ps^r$ as $s \downarrow 0$. From (5.17) we have

$$\begin{aligned}\psi''(x) = (1-x)^{-1}\left(\frac{2\rho(x)\rho''(x)}{\rho'(x)^2} - 1\right) - (1-x)^{-2}\frac{\rho(x)}{\rho'(x)} \\ - \left(\frac{\rho''(x)}{\rho'(x)} + \frac{\rho(x)\rho'''(x)}{\rho'(x)^2} - \frac{2\rho(x)\rho''(x)^2}{\rho'(x)^3}\right)\log(1-x).\end{aligned} \tag{5.18}$$

Taking $x \downarrow 0$ in (5.17) and (5.18), using $\rho^{(k)}(x) \sim \frac{r!}{(r-k)!}px^{r-k}$ for $k \leq 3$, we obtain

$$\psi'(0) = 0; \quad \psi''(0) = \frac{r-2}{r} > 0,$$

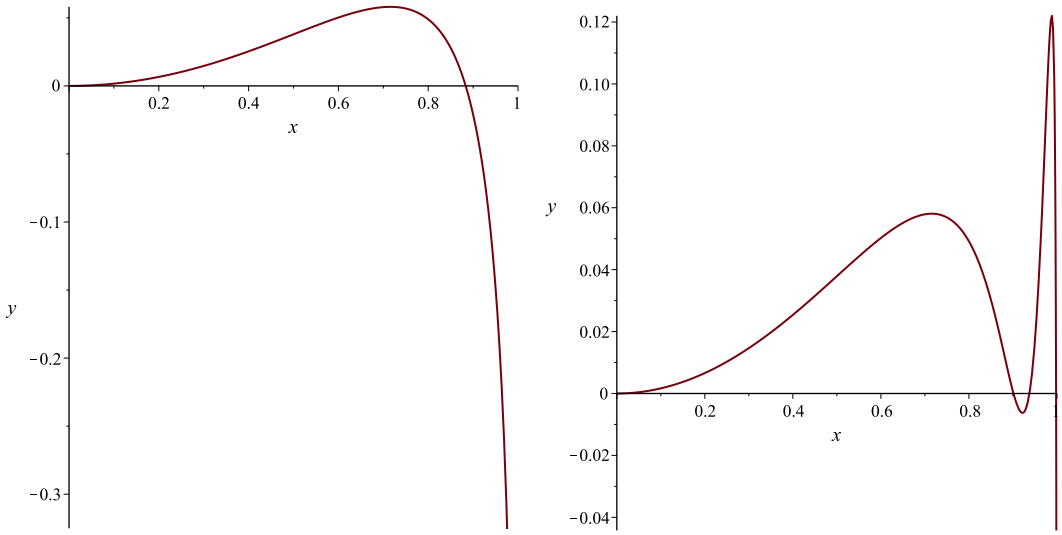

Figure 3: Plots of $y = \psi(x)$ for $\rho(s) = s^3$ (*left*) and $\rho(s) = 0.9s^3 + 0.1s^{38}$ (*right*). On the left, the only positive zero is $x_1 \approx 0.883414$. On the right, the 3 positive zeros are $x_1 \approx 0.901174$, $x_2 \approx 0.937414$, and $x_3 \approx 0.997979$. On the right, $\alpha_\rho^\sharp \approx 0.872923$ and $g^*(\alpha_\rho^\sharp) \approx 0.988192$, and only the zero x_3 exceeds this value. Proposition 5.7 gives $\underline{\alpha}_\rho \approx 0.917935$ for the case on the left and $\underline{\alpha}_\rho \approx 0.998263$ for the case on the right.

since $r \geq 3$. Hence $\psi(0) = 0$ is a local minimum, and $\psi(x) > 0$ for $x > 0$ small enough. But $\psi(x) \to -\infty$ as $x \uparrow 1$, so continuity implies that ψ has at least one zero in $(0, 1)$.

Consider the condition (a) in the proposition. Suppose that h has a unique local minimum located at $x_\rho^\sharp \in (0, 1)$, so $\alpha_\rho^\sharp = h(x_\rho^\sharp)$. Then by Lemma 5.8, ψ has a unique local maximum at $x_\rho^\sharp$, and necessarily $\psi(x_\rho^\sharp) > 0$. By continuity (and Rolle's theorem) it follows that ψ has exactly one zero $x_\rho^* \in (x_\rho^\sharp, 1)$. So (a) $\Rightarrow$ (b). Moreover, it follows that $\psi(x) > 0$ for $x \in (0, x_\rho^*)$ and $\psi(x) < 0$ for $x > x_\rho^*$. The claim (5.15) follows.

Finally, suppose $\rho(s) = s^r$ for some $r \geq 3$. We deduce (a). First note (cf (5.18))

$$\psi''(x) = \frac{1}{r(1-x)} \left(r - 1 - \frac{1}{1-x} \right).$$

Hence $\psi''(0) = \frac{r-2}{r} > 0$ and for $x \in (0, 1)$ we have $\psi''(x) = 0$ if and only if $x = \frac{r-2}{r-1}$ (an inflexion point). By Rolle's theorem and the fact that $\psi(x) \to -\infty$ as $x \uparrow 1$, $\psi'(x) = 0$ at exactly one $x \in (0, 1)$, necessarily a local maximum $x \in (\frac{r-2}{r-1}, 1)$. Another application of Rolle's theorem shows that ψ has a single positive zero, necessarily in $(\frac{r-2}{r-1}, 1)$. □

Now we can complete the proof of Theorem 2.2.

Proof of Theorem 2.2. Markov's inequality applied to $\mathcal{N}(n, m) - 1 \geq 0$ gives

$$\mathbb{P}_{\rho_n}[T_n \leq m] = \mathbb{P}_{\rho_n}[\mathcal{N}(n, m) \geq 2] \leq \mathbb{E}_{\rho_n}[\mathcal{N}(n, m)] - 1. \tag{5.19}$$

Suppose that $m_n/n \to \alpha \in (0, \alpha_\rho^*)$. Then by (5.19) with (2.4), $\mathbb{P}_{\rho_n}[T_n \leq m_n] = O(n^{-1})$. It follows that, for any $\varepsilon > 0$, $\mathbb{P}_{\rho_n}[T_n \leq (\alpha_\rho^* - \varepsilon)n] \to 0$. On the other hand, Theorem 2.4(iii) implies that there exists $\delta > 0$ such that for any $\alpha \in (\underline{\alpha}_\rho, \underline{\alpha}_\rho + \delta)$, $\mathbb{P}_{\rho_n}[T_n \leq \alpha] \to 1$. Moreover, these results together show that $\alpha_\rho^* \leq \underline{\alpha}_\rho$, and $\underline{\alpha}_\rho \leq 1$ by Proposition 5.6. □

A Threshold numerics and asymptotics

In this appendix we return to the discussion of the fixed-weight case presented in Section 2.5. By (2.1) we have $\alpha_r^* := \inf\{\alpha \geq 0 : F_r(\alpha) > 0\}$, where

$$F_r(\alpha) := \log \sup_{\gamma \in [0,1/2]} \left(\frac{(1+(1-2\gamma)^r)^\alpha}{2\gamma^\gamma (1-\gamma)^{1-\gamma}} \right). \tag{A.1}$$

If $r \geq 3$, ψ has a single positive zero in $x_r^* \in (0,1)$ (see Proposition 5.7) satisfying

$$x_r^* = -\left(1 - \left(\frac{r-1}{r}\right) x_r^*\right) \log(1 - x_r^*), \tag{A.2}$$

$$\text{and } \underline{\alpha}_r = h(x_r^*) = -\frac{\log(1-x_r^*)}{r(x_r^*)^{r-1}}. \tag{A.3}$$

For example, $\alpha_3^\sharp \approx 0.818469$, $g^*(\alpha_3^\sharp) \approx 0.715332$, and $x_3^* \approx 0.883414$ so $\underline{\alpha}_3 \approx 0.917935$ (see also Figure 3).

We consider the evaluation (and asymptotics) of $\underline{\alpha}_r$; this will lead to a proof of Proposition 2.7. One can obtain arbitrarily sharp upper and lower bounds for x_r^* as follows. By (A.2) we have that $x_r^* = i_r(x_r^*)$, where we set

$$i_r(x) := 1 - \exp\left\{ -\frac{x}{1 - (\frac{r-1}{r})x} \right\}.$$

For $\theta \in [0,1)$, $x \mapsto \frac{x}{1-\theta x}$ is strictly increasing for $x \in [0,1]$. Thus if $x_r^* > a_n$, it follows that $x_r^* > a_{n+1} := i_r(a_n)$. Also, $i_r'(0) = 1$, $i_r''(0) = \frac{2-r}{r} > 0$, and $i_r(1) = 1 - \mathrm{e}^{-r} < 1$, so $i_r(x) > x$ for $x \in (0, x_r^*)$ but $i_r(x) < x$ for $x \in (x_r^*, 1]$. Hence starting with $a_0 = \frac{r-2}{r-1} < x_r^*$ (an inequality in Proposition 5.7), iteration yields an increasing sequence of lower bounds a_n for x_r^*. Conversely, starting with $b_0 = 1 > x_r^*$ and iterating $b_{n+1} := i_r(b_n)$ gives a decreasing sequence of upper bounds b_n for x_r^*. For example, after one step we get

$$1 - \exp\left\{ -\frac{r(r-2)}{2(r-1)} \right\} < x_r^* < 1 - \mathrm{e}^{-r}, \quad (r \geq 3).$$

Proceeding up to b_2 for the upper bound and a_4 for the lower bound is sufficient to obtain the $r \to \infty$ asymptotic expression

$$x_r^* = 1 - \mathrm{e}^{-r} - r^2 \mathrm{e}^{-2r} + O(r^4 \mathrm{e}^{-3r}), \tag{A.4}$$

which will be the main ingredient in the proof of the $\underline{\alpha}_r$ result in (2.17).

In fact, this iteration converges, so $a_n \uparrow x_r^*$ and $b_n \downarrow x_r^*$. To prove convergence it suffices to show that $i_r'(x) < 1$ at $x = x_r^*$. A calculation shows that $i_r'(x)$ evaluated at $x = x_r^*$ is $x^{-2}(1-x)(\log(1-x))^2$, so for the required inequality it suffices to show that

$$-x^{-1}\log(1-x) < (1-x)^{-1/2}, \text{ for } 0 < x < 1.$$

The coefficient of x^k in the power series expansion of the left-hand side of the last inequality is $1/(k+1)$, and for the right-hand side it is $4^{-k}\binom{2k}{k}$, and both series are convergent on the given interval. An induction shows that $1/(k+1) \leq 4^{-k}\binom{2k}{k}$ for all integers $k \geq 0$. So term-by-term comparison of the two power series gives the inequality.

Now we present the proof of Proposition 2.7.

Proof of Proposition 2.7. Take $\rho(s) = s^r$ for $r \geq 3$. As already mentioned, the asymptotic for α_r^* is in [3]. Let $\alpha > 0$. Then, by (2.9),

$$h(1 - \mathrm{e}^{-\alpha r/2}) = -\frac{\log(\mathrm{e}^{-\alpha r/2})}{r(1-\mathrm{e}^{-\alpha r/2})^{r-1}} = \frac{\alpha}{2}(1 + o(1)),$$

as $r \to \infty$. Hence $h(1-\mathrm{e}^{-\alpha r/2}) \le \alpha$ for all r sufficiently large, which by (2.10) shows that $\limsup_{r\to\infty} \alpha_r^\sharp \le \alpha$. Since $\alpha > 0$ was arbitrary, it follows that $\lim_{r\to\infty} \alpha_r^\sharp = 0$.

Finally, by (A.3) with (A.4) and repeated Taylor expansions we obtain $\underline{\alpha}_r = 1 - \mathrm{e}^{-r} + O(r^3\mathrm{e}^{-2r})$, completing the proof of (2.17). □

We briefly review previous descriptions of α_r^* in the literature. Calkin [3, §4] gives the same description of α_r^* as our (A.1). Systems of nonlinear equations for α_r^* are proposed in [2, 15, 4]; these can be shown to be consistent with our description, as follows. Let

$$F_{r,\alpha}(\gamma) := \log\left(\frac{(1+(1-2\gamma)^r)^\alpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}\right), \tag{A.5}$$

so that $F_r(\alpha) = \sup_{\gamma\in[0,1/2]} F_{r,\alpha}(\gamma)$. Differentiating, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}F_{r,\alpha}(\gamma) = -\frac{2\alpha r(1-2\gamma)^{r-1}}{1+(1-2\gamma)^r} + \log\left(\frac{1-\gamma}{\gamma}\right). \tag{A.6}$$

Then one may characterize α_r^* by the two equations $F_{r,\alpha}(\gamma) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}\gamma}F_{r,\alpha}(\gamma) = 0$. On the substitution $\lambda = \frac{1}{2}\log(\frac{1-\gamma}{\gamma})$, the first of these equations becomes, after some calculations along the lines of those in Section 3.5,

$$(1+(\tanh\lambda)^r)^\alpha \mathrm{e}^{-\lambda\tanh\lambda}\cosh\lambda = 1. \tag{A.7}$$

For the second equation, setting (A.6) to zero gives, after the same substitution for λ,

$$r\alpha = (1+(\tanh\lambda)^{-r})\lambda\tanh\lambda. \tag{A.8}$$

The system defined by equations (A.7) and (A.8) is the same as that given by Cooper [4, p. 269], and, after some manipulation, is seen to coincide also with that given by Balakin *et al.* [2, p. 564] and Kolchin [15, p. 139]. Small discrepancies in the values for α_r^* given in [2, 15, 3, 4] can presumably be put down to numerical inaccuracies.

Finally, we discuss the numerical evaluations of α_r^* in Table 1; here we use some claimed properties of the functions involved that we do not verify rigorously. Observe from (A.6) that any stationary value γ for $F_{r,\alpha}$ solves

$$\alpha = \frac{1+(1-2\gamma)^r}{2r(1-2\gamma)^{r-1}}\log\left(\frac{1-\gamma}{\gamma}\right) =: \alpha_r(\gamma). \tag{A.9}$$

Numerical curve sketching shows that (A.9) generically has at most 2 solutions in $(0,1/2)$; of such solutions, the smallest will be the local maximum, since $F'_{r,\alpha}(\gamma) \to \infty$ as $\gamma \downarrow 0$, by (A.6). If (A.9) has no solutions in $(0,1/2)$, then the supremum in (A.1) is either $F_{r,\alpha}(0) = (\alpha-1)\log 2$ or $F_{r,\alpha}(1/2) = 0$. Thus setting $\gamma_0 := \gamma_0(\alpha,r) = 0$ if (A.9) has no solutions in $(0,1/2)$ and $\gamma_0 := \gamma_0(\alpha,r)$ to be the smallest positive solution to (A.9) otherwise, we have that $F_r(\alpha) = F_{r,\alpha}(\gamma_0)$ whenever $\alpha \in (0,1)$.

For $\alpha \in (0,1)$, $F_r(\alpha) > 0$ if and only if $\gamma_0(\alpha,r) > 0$. Moreover, Lemma 4.1 shows that $\alpha_r^* < 1$, so that for $\alpha < 1$ such that $\gamma_0(\alpha,r) > 0$, $F_r(\alpha) = \alpha_r(\gamma_0)\log(1+(1-2\gamma_0)^r) - \log(2\gamma_0^{\gamma_0}(1-\gamma_0)^{1-\gamma_0})$. Thus to find α_r^*, we solve for $\gamma \in [0,1/2]$ the equation

$$\alpha_r(\gamma) - \phi_r(\gamma) = 0, \tag{A.10}$$

where

$$\phi_r(\gamma) = \frac{\log\left(2\gamma^\gamma(1-\gamma)^{1-\gamma}\right)}{\log\left(1+(1-2\gamma)^r\right)}.$$

Numerical curve plotting shows that $\gamma \mapsto \alpha_r(\gamma) - \phi_r(\gamma)$ is decreasing on $[0,1/2]$, so (A.10) can be solved using efficient numerical methods; let γ_r denote the solution to (A.10). Then we compute α_r^* via $\alpha_r^* = \alpha_r(\gamma_r)$.

B Technical appendix

This appendix gives two technical results that we need in the body of the paper. The first, whose proof is omitted, collects some elementary properties of probability generating functions (see e.g. [12, pp. 264–266]).

Lemma B.1. *Let $\phi(s) := \mathbb{E}[s^X]$, $s \in [-1,1]$, for a $\mathbb{Z}_+$-valued random variable X. Then $\phi(0) = \mathbb{P}[X=0]$, $\phi(1) = 1$, and $\phi(s)$ is infinitely differentiable at least for $s \in (-1,1)$; if $\mathbb{E}[X] < \infty$ then $\phi'(s) = \frac{\mathrm{d}}{\mathrm{d}s}\phi(s)$ is continuous in the closed interval $[-1,1]$. Moreover:*

(i) Suppose that $\mathbb{P}[X=0] = 0$. Then as $s \downarrow 0$,

$$\phi(s) = s\mathbb{P}[X=1] + O(s^2), \text{ and } \phi'(s) = \mathbb{P}[X=1] + O(s).$$

(ii) If $\mathbb{E}[X] < \infty$, then as $s \downarrow 0$,

$$\phi(1-s) = 1 - s\mathbb{E}[X] + o(s), \text{ and } \phi'(1-s) = \mathbb{E}[X] + o(1).$$

(iii) For any $s \in [0,1]$, $|\phi(-s)| \le \phi(s)$.

We shall use the following bounds on the binomial coefficient $\binom{n}{k}$.

Lemma B.2. *Let $n \in \mathbb{N}$ and $k \in \{0,1,\dots,n\}$. Then*

$$\binom{n}{k} \le \left(\left(\frac{k}{n}\right)^{k/n} \left(1-\frac{k}{n}\right)^{1-(k/n)} \right)^{-n} \le n^k \mathrm{e}^k k^{-k}. \tag{B.1}$$

On the other hand, if $0 < k < n$,

$$\binom{n}{k} \ge \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \mathrm{e}^{-1/6} \left(\left(\frac{k}{n}\right)^{k/n} \left(\frac{n-k}{n}\right)^{(n-k)/n} \right)^{-n}. \tag{B.2}$$

Proof. Robbins's refinement of Stirling's formula (see e.g. [12, §II.9]), says that

$$n! = (2\pi)^{1/2} n^{n+(1/2)} \mathrm{e}^{-n+\varepsilon_n}, \text{ for any } n \ge 1,$$

where $\frac{1}{12n+1} < \varepsilon_n < \frac{1}{12n}$. This yields the upper bound, for $n \ge 1$ and $k, n-k \ge 1$,

$$\binom{n}{k} \le \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \left(\left(\frac{k}{n}\right)^{k/n} \left(\frac{n-k}{n}\right)^{(n-k)/n} \right)^{-n}, \tag{B.3}$$

where we have used the fact that

$$\varepsilon_n - \varepsilon_k - \varepsilon_{n-k} \le \frac{1}{12n} - \frac{12n+2}{144k(n-k)+12n+1} \le \frac{1}{12n} - \frac{12n+2}{36n^2+12n+1} \le 0.$$

Considering separately the cases (i) $k \in \{0,n\}$, and (ii) $0 < k < n$, using (B.3) in case (ii), we obtain the first inequality in (B.1). The second follows from the fact that

$$\left(1-\frac{k}{n}\right)^{-(n-k)} = \left(1+\frac{k}{n-k}\right)^{n-k} \le \mathrm{e}^k.$$

For the lower bound, another application of Robbins's bounds yields (B.2), where for the $\mathrm{e}^{-1/6}$ term we have used the fact that $\varepsilon_n - \varepsilon_k - \varepsilon_{n-k} \ge -\frac{1}{12k} - \frac{1}{12(n-k)} \ge -\frac{1}{6}$. □

References

[1] Alamino, R.C. and Saad, D.: Typical kernel size and number of sparse random matrices over Galois fields: A statistical physics approach. *Phys. Rev. E* **77**, (2008), 061123. MR-2496138

[2] Balakin, G.V., Kolchin, V.F. and Khokhlov, V.I.: Hypercycles in a random hypergraph. *Discrete Math. Appl.* **2**, (1992), 563–570. MR-1138095

[3] Calkin, N.J.: Dependent sets of constant weight binary vectors. *Combinat. Probab. Comput.* **6**, (1997), 263–271. MR-1464565

[4] Cooper, C.: Asymptotics for dependent sums of random vectors. *Random Struct. Alg.* **14**, (1999), 267–292. MR-1680224

[5] Cooper, C.: The cores of random hypergraphs with a given degree sequence. *Random Struct. Alg.* **25**, (2004), 353–375. MR-2099209

[6] Costello, K.P. and Vu, V.: On the rank of random sparse matrices. *Combinat. Probab. Comput.* **19**, (2010), 321–342. MR-2607371

[7] Darling, R.W.R., Levin, D.A. and Norris, J.R.: Continuous and discontinuous phase transitions in hypergraph processes. *Random Struct. Alg.* **24**, (2004), 397–419. MR-2060628

[8] Darling, R.W.R. and Norris, J.R.: Structure of large random hypergraphs. *Ann. Appl. Probab.* **15**, (2005), 125–152. MR-2115039

[9] Darling, R.W.R. and Norris, J.R.: Differential equation approximations for Markov chains. *Probab. Surv.* **5**, (2008), 37–79. MR-2395153

[10] Dietzfelbinger, M., Goerdt, A., Mitzenmacher, M., Montanari, A., Pagh, R. and Rink, M.: Tight thresholds for cuckoo hashing via XORSAT. *Proc. 37th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, 2010, Volume 6198/2010, Springer, pp. 213–225.

[11] Dubois, O. and Mandler, J.: The 3-XORSAT threshold. *Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002, pp. 769–778 (version dated 28 February 2003).

[12] Feller, W.: An Introduction to Probability Theory and Its Applications, Vol. I. 3rd ed., John Wiley, New York, 1968. MR-0228020

[13] Ibrahimi, M., Kanoria, Y., Kraning, M. and Montanari, A.: The set of solutions of random XORSAT formulae. *SODA '12 Proc. 23rd Annual ACM-SIAM Symposium on Discrete Algorithms*, 2012, pp. 760–779.

[14] Kolchin, V.F.: Cycles in random graphs and hypergraphs (abstract). *Adv. in Appl. Probab.* **24**, (1992), 768.

[15] Kolchin, V.F.: Random graphs and systems of linear equations in finite fields. *Random Struct. Alg.* **5**, (1994), 135–146. MR-1248181

[16] Kolchin, V.F.: Random Graphs. Cambridge University Press, 1999. MR-1728076

[17] Kolchin, V.F., Sevastyanov, B.A. and Chistyakov, V.P.: Random Allocations. V.H. Winston & Sons, Washington, 1978. MR-0471016

[18] Kovalenko, I.N. and Levitskaya, A.A.: Stochastic properties of systems of random linear equations over finite algebraic structures. Probabilistic Methods in Discrete Mathematics, V.F. Kolchin *et al.* (eds.), TVP/VSP, Utrecht, 1993, pp. 64–70. MR-1383129

[19] Levitskaya, A.A.: Systems of random equations over finite algebraic structures. *Cybernet. Systems Anal.* **41**, (2005), 67–93. MR-2188375

[20] Mahmoud, H.M.: Pólya Urn Models. CRC Press, 2009. MR-2435823

[21] Mézard, M., Parisi, G. and Zecchina, R.: Analytic and algorithmic solution of random satisfiability problems. *Science* **297**, (2002), 812–815.

[22] Muetze, T.: Generalized switch-setting problems. *Discrete Math.* **307**, (2007), 2755–2770. MR-2362960

[23] Talagrand, M.: Spin Glasses: A Challenge for Mathematicians. Springer-Verlag, Berlin, 2003. MR-1993891

Acknowledgments. The authors thank Julian West for pointing out the elementary application of the pigeonhole principle in the proof of Lemma 5.1, thus avoiding the use of linear algebra, and the anonymous referees for their comments and suggestions on the first version of this paper, which have led to significant improvements in the presentation.